Musculoskeletal pain & dysfunction in musicians

Woldendorp, Kees Hein

Chapter 5

Clinical assessment of global posture
in routine practise

Submitted as:
Kees H. Woldendorp, Jonas F.E. Kleinbergen, Anne M. Boonstra,
Antoine W. de Schipper,
J. Hans Arendzen, Michiel F. Reneman.
Quality and usability of clinical assessments of static standing and sitting posture in
musicians: a systematic review.
(Submitted in 2018).

Quality and usability of clinical assessments of static standing and sitting posture in musicians: a systematic review

Abstract

Background:

A validated method to assess sitting and standing posture in a clinical setting is needed to guide diagnosis, treatment and evaluation of these postures in musicians. At present, no systematic overview of assessment methods, their clinimetric properties, and usability is available. The objective of this study was to provide such an overview and to interpret the results for clinical practice.

Methods:

Systematic literature review according to international guidelines. Two independent reviewers assessed risk of bias and clinimetric values of the assessment methods, and their usability (using a self-constructed checklist). Quality of evidence and strength of recommendations were determined according to the Grading of Recommendations Assessment, Development and Evaluation working group (GRADE).

Results:

Out of 27,680 records, 41 eligible studies were included. Thirty-two assessment instruments were identified, clustered into five categories: direct & indirect visual observation, direct & indirect body measurement, and digital assessment. The methodological quality of 27 (66%) of the studies was moderate to good. Reliability was most frequently studied. Little information was found about validity and none about responsiveness.

Conclusions:

Based on a moderate level of evidence, a tentative recommendation can be made to use a direct or indirect visual observation method with global posture recorded by a trained observer as a rating scale.

Keywords: measurement - analysis – observation – musculoskeletal disorders – reliability – validation

Introduction

Among musicians there is a high prevalence of musculoskeletal complaints [1]. A causal relation is often assumed between ‘poor’ postures and musculoskeletal complaints in both musicians and non-musicians [2-6]. Identification of asymmetries and other ‘abnormalities’ during static posture is a common procedure in the clinical practice of music medicine, physical therapy, rehabilitation medicine and occupational medicine [7-8]. It is not clear what a ‘poor’ or ‘risky’ posture may be [9-12], nor is there agreement about ways to perform and record observations of sitting and standing postures with a valid, reliable, and clinically usable method.

Reliable information about (working) posture is important in the diagnosis and treatment of patients with musculoskeletal complaints, to detect potentially risky postures, under the assumption that changing these postures will decrease the complaints [3,13-16]. In addition, evaluating the results of therapy and comparing the effects of different therapies to improve posture requires an assessment method that is sensitive to posture changes.

Despite a wide range of literature about aspects of posture assessment, there is little literature about the clinimetric aspects of the methods used for this in daily practice. This is the case for musicians, but also for non-musicians. As far as we are aware, there have been few systematic reviews performed in accordance with the international guidelines and focusing on assessment methods for global posture – as opposed to specific aspects of posture – that might be suitable for any standing or sitting patient (including musicians). Musicians are singled out here as they are a subgroup of patients with a high prevalence of musculoskeletal complaints [1], and therefore of special interest in clinical practice.

The most valid and reliable assessment methods, such as multi-camera systems like Optotrak, Vicon, Motion Analysis, or Surface Topography Systems [17-20], are expensive and time-consuming, making large-scale use of this kind of instrumentation in routine clinical settings unrealistic. On the other hand, the widely used assessment method for posture in routine practice, i.e. visual observation by a clinician, seems to have low intra- and inter-observer reliability [7,21]. Moreover, visual inspection is normally not performed in a standardized way. Although training of observers appears to improve the levels of agreement, they are still moderate [20,22].
In order to find a clinically useful and reliable method to assess posture, especially one which can be used in the treatment of juvenile and adult musicians, we performed a systematic review with the intention to identify a clinically useful and reliable method to assess static posture. The aim of this study was to provide an overview of the clinimetric and feasibility properties of the assessment methods for static standing and/or sitting posture in a routine clinical setting, and to interpret the findings for clinical practice. In view of the limited number of publications focusing on musicians, we have widened the scope of our review to include posture assessment of all kinds of sub-populations in clinical practice.

Methods

Operationalization of the research objective
The terms used in describing the aim of the study were defined as follows:

- 'Assessment method for posture' includes all types of standardized methods by which the posture of a human being can be assessed visually or with the help of, e.g., photography.

- 'Clinimetric properties' (including interpretation, recording, and evaluation) can be assessed in qualitative ways (e.g. 'good/not good', 'risky posture' or 'better/worse') and/or quantitative ways (e.g. 'millimeters/degrees', 'data plotted against reference data for a population' or 'difference in millimeters/degrees').

- 'Suitable for routine practice in a normal clinical setting' means that the instrument is inexpensive, not too space-consuming, transportable and easy to use without extensive training. Similar requirements apply to the technical aspects. It is important, for example, that the data obtained should be delivered to health care professionals such as physical therapists, ergonomists, and/or physicians in a simple format and without delays.

- 'Posture' is the alignment or orientation of the body segments while maintaining a position [23].
  - o 'Static' means that the aspects of movement, maintaining balance or other time-related dynamics are not included.
  - o 'Sitting posture', in the absence of an internationally agreed scientific definition [24], we define this as the situation in which the body is resting on a seat on the buttocks or haunches [25].
  - o 'Standing posture' is the position in which a person stands upright with at least one foot on the ground for more than 4 seconds, while remaining within a 1 m^2 area [26].

Search
First, electronic medical databases, one trial register, and additional non-electronic channels (grey literature) were searched for eligible articles. The database search was conducted on December 1, 2017, in accordance with the Cochrane guidelines for systematic reviews of diagnostic tests [27]. It covered the electronic databases Cochrane (1940-2017), Medline (PubMed) (1950-2017), Embase (1974-2017), CISDOC (1901-2017), ScienceDirect (1997-2017), Web of Science (1900-2017) and CINAHL (1977-2017). An additional search (using the search terms 'posture AND assessment') was performed in ClinicalTrials.gov in December 2017. 'Grey literature' was searched from December 1, 2014 through December 31, 2017.

Search terms (MeSH and non-MeSH terms) were divided into three domains: 'the instrument' (e.g. method, instrument, technique); 'the goal of the instrument' (e.g. assessing, screening, examining); and 'posture' (e.g. upright position, posture, seated position). Individual search terms within each of the domains were combined with the Boolean operator 'OR'. The three domains themselves were combined with the Boolean operator 'AND'. It was anticipated that a huge amount of records would emerge from the databases, in view of the broad scope of the search terms, and since this is a common feature of systematic reviews about measurement properties [28]. Therefore, in order to keep the number of records manageable, we added a fourth domain, linked to the other three by the Boolean operator 'NOT', to exclude non-relevant titles. Search items were added to this fourth domain until the number of records in the first database (Medline) had dropped to below 15,000 records (Supplementary table S1). In the procedure to reduce the number of titles by using 'NOT' + MeSH-terms, and in order not to lose any potentially relevant records, we checked the validity of the procedure by checking whether three of the very relevant records [11,22,29] found in a previous explorative review (performed by our group) remained in the pool of records.
The search strategies differed slightly for each database (Supplementary table S1).

The references of the relevant papers were searched, as were the reference lists of papers thus identified, and the reference lists of 13 identified review papers about posture assessment [8,11,22,30-39].

'Grey' literature was collected through various non-electronic channels, i.e. via colleagues, from books, and by means of a hand search of the journal Medical Problems of Performing Artists from 1986 to 2000. Finally, duplicate articles were removed.

Selection

The selection procedure is presented in Supplementary figure S1. Titles, abstracts and full texts were screened independently by two reviewers (KHW and JK), in three stages, for their eligibility according to the inclusion and exclusion criteria. This resulted in three pools of potentially relevant records (pools 2, 3, and 4a), as shown in table 1. Additional exclusion criteria were added after pool 4A had been created, because the number of articles was still too large: we excluded articles in other languages than English, articles about assessment instruments limited to the observation of <2 adjacent domains of body posture(s) (e.g. back and lower extremities), and articles published before 1990. This resulted in the final pool 4b (see table 1 for inclusion and exclusion criteria).

Table 1: In- and exclusion criteria.

	Inclusion criteria (Pool 1-4A) #		**Exclusion criteria** (Pool 4A)#
1	All languages,	1	Articles about assessment instruments for the range of movement or movement,
2	Articles about assessment instruments of the observation of ≥ 1 domain of body posture(s),	2	Articles about assessment instruments for body balance,
3	Articles about assessment instruments and the assessment method of posture.	3	Articles about assessment instruments using a (skills) lab or other complex /expensive/time spending method not applicable for daily use,
4	Articles about assessment instruments based on validation of the instrument (level of evidence A2, B or C *).	4	Articles about assessment instruments measuring over a period of team, with e.g. the mean or number of posture frequencies over time as outcome,
		5	Articles about assessment instruments based on interpretation by the authors (systematic review or experts opinion: level of evidence A1 or D *),
		6	Articles about assessment instruments that provided insufficient Information to allow adequate interpretation of outcome measures and results.
			Additional exclusion criteria (Pool 4B)#
		7	Non-English papers,
		8	Papers published before 1990,
		9	Articles about assessment instruments of the observation of < 2 abutting domains of body posture(s) (e.g. head and lower extremities),

 # Pool 4A of records: the pool of potential relevant records in the initial search, Pool 4B: final pool of included records, created by additional exclusion criteria for reasons of handling (see text).
** CBO-Levels of Evidence (2007): see www.cbo.nl for detailed Information.*

At each stage of screening (title, abstract, and full text), the reviewers (KHW and JK) met and resolved disagreement about individual citations through consensus or, if necessary, by consulting a third reviewer (AMB). The two reviewers merged the data into one database, and checked whether all data had been entered correctly. The same procedure was performed for entering data in the final tables and figures. The level of agreement between the two reviewers was calculated at all stages using % agreement and Cohen's kappa.

Missing information
If papers about the clinimetric values of an assessment instrument referred to other publications about the development of that instrument, we included these papers as part of the primary paper. If data extraction was not possible, additional information was obtained by contacting the authors listed in the paper. Missing information was recorded in the final critical appraisal tables.

Assessment of quality
The methodological quality of each included study was assessed independently by the two researchers using the Quality Assessment of Diagnostic Accuracy Studies-II (QUADAS-II) checklist [40] and the COnsensus-based Standards for the Selection of health Measurement INstruments (COSMIN) [41,42]. The combined use of these two checklists provided complementary information, despite some overlap: items 3, 4, and 14 in QUADAS-II are identical to H4, B7 and F1, F2, H5, respectively, in COSMIN. Most items in COSMIN that refer to the presence of restrictions regarding design requirements and statistical methods are asked for more specifically in the QUADAS-II items. Items 1, 2 and 6-13 of the QUADAS-II are not included in COSMIN.
The QUADAS-II instrument consists of four domains: patient selection, index test, reference standard, and 'flow and timing' (flow of patients through the study and timing of the index tests and reference standard) [40]. We graded the risk of bias in patient selection, index test, reference standard, and flow and timing as high, unclear, or low. The same assessment strategy was used for applicability regarding patient selection, index test, and reference standard.

The items (boxes) of COSMIN [41,42] were used to determine the clinimetric values of the instruments. Interpretation of the correlation coefficient of reliability was as follows: values >0.75 were regarded as good, those between 0.50 and 0.75 as moderate and those <0.5 as poor reliability [43]. Interpretation of the correlation coefficient for criterion and concurrent validity was as follows: values ≥ |0.70 |were regarded as strong, between |0.3 | and |0.70 |as moderate and values ≤ |0.3 |as weak [43].

A self-constructed customized checklist (Supplementary table S2) was developed to measure the clinical usability of each measurement instrument for posture. Aspects included in the checklist were readability of the instructions, comprehensibility, time required to administer the instrument, physical requirements (e.g. camera, space, researcher) and the effort involved in interpretation. Each aspect was scored with points ranging from -2 to 2 or -1 to 1. A sum score for each posture measurement instrument was calculated by summing the item scores. Sum scores were calculated for both the clinimetric aspects and clinical usability to enable us to balance the clinical use and scientific support for each measurement instrument.

Interpretation of results

Results were aggregated and interpreted according to the framework for therapeutic and diagnostic tests developed by the Grading of Recommendations Assessment, Development and Evaluation (GRADE) working group [44-49]. The framework for therapeutic studies covers five aspects of quality of evidence (study design, inconsistency of results, indirectness of evidence, imprecision, and reporting bias) and four aspects of the strength of recommendation (quality of the evidence, uncertainty about the balance between desirable and undesirable effects, uncertainty or variability in values and preferences, and uncertainty about whether the intervention involves extensive use of resources) [45].

The above five aspects of quality of evidence for therapeutic studies are categorized in a somewhat different way for diagnostic studies; in addition to the items of study design and limitations/risk of bias, the other aspects are all allocated to the item of indirectness of evidence (outcomes, patient populations, diagnostic test, comparison test, and indirect comparisons, inconsistency of outcomes, imprecise evidence, and publication bias). We adapted the GRADE framework as follows to make it suitable for our study. While the GRADE framework for diagnostic studies rates cross-sectional study designs as high quality of evidence (initial score of 4) [48], we applied the same score to observational studies. Risk of bias and patient population, including concerns about applicability, were determined with QUADAS-II. Inconsistency of outcomes and imprecise results were rated for each study, using the inter-rater reliability values in two ways: values of outcomes were determined by the scores themselves, while inconsistency of outcomes was based on the difference between the maximum and minimum values for each study (see below). Imprecision of results was evaluated as measurement error. The aspect of cost was taken into account in the usability item.

Missing data were not taken into account in the aggregated score for a group of assessment methods. At the level of types of measurement instrument, each aspect was scored by taking the rounded mean value for that aspect for all the relevant studies

Points were added to or deducted from the initial score as follows:

- Interpretation of risk of bias and patient population (concerns about applicability): This was based on the QUADAS-II scores: 0 = ≤1 uncertain score/study; -1 = ≤3 uncertain scores or ≤1 high score/study; -2 = >4 uncertain scores or >1 high score/study.

- Scores for outcomes:
  These were based on the values of inter-rater reliability:
  - Values of outcomes: The following cut-off points were chosen, partly in line with Portney and Watkins [43] and partly arbitrarily: Overall scores: 1 = if ≥75% of

correlation coefficients exceed 0.75/study; 0 = if ≥75% of correlations coefficients exceed 0.5/study, but not score 1; -1 = other;
- Inconsistency of outcomes: 0 = consistent outcomes per study (range of outcomes is small, within the width of one category of correlation coefficient scores) [42]; -1 = much variability among outcomes per study (range of outcomes greater than one category of correlation coefficient scores);

- Imprecision of results: The following cut-off points were chosen arbitrarily: Measurement error: 0 = ≤10 mm or ≤10 degrees; -1 = >10 mm or >10 degrees.

- Usability scores (Supplement 3): The following cut-off points were chosen arbitrarily: 1 = ≥9 points/study; 0 = 3 ≤ points/study < 9; -1 = < 3 points/study.

The probability of publication bias was determined by comparing the size of the study sample with the level of the inter-rater reliability values. If smaller studies (<30 participants) had higher inter-rater reliability values than the larger studies (≥30 participants), this could be an indication of publication bias.

The strength of recommendations was rated as follows: A (sum score 4) = strong recommendation to use this method, B (sum score 3) = weak recommendation to use this method, C (sum score 2) = weak recommendation not to use this method, D (sum score 1) = strong recommendation not to use this method in clinical practice [45].

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist of 'items to include when reporting a systematic review' [50] and the checklist 'A Measurement Tool to Assess Systematic Reviews' (AMSTAR) [51] were used to optimize our reporting of the present review. The study protocol was accepted for registration in the PROSPERO register (no. CRD42017041711).

Results

Search

The first search identified 27,680 papers, 389 of which were retrieved in full text and screened for eligibility. In the end, 41 of these papers were included in the review. Results of the screening and selection process are presented in figure 1. For one record, [21] the decision to include was made by the third reviewer (AMB). In the final step of the selection process, the agreement between the two screeners was good (K= 0.66) (figure 1).

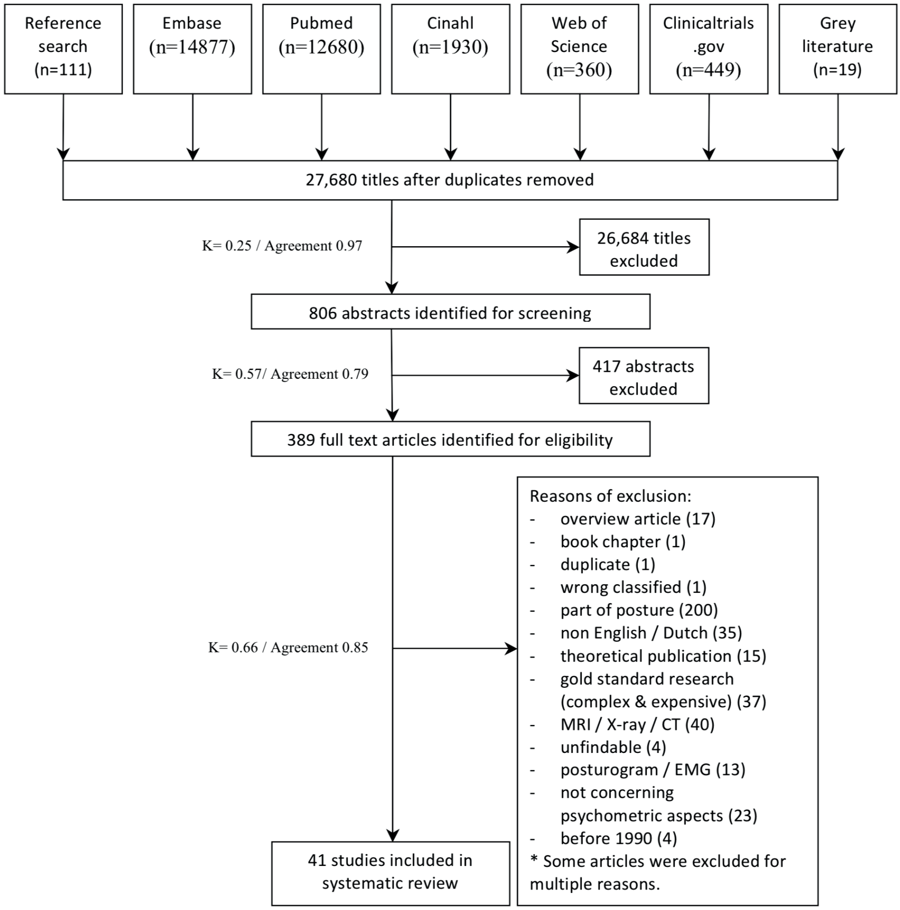

Figure 1: Flow chart of the screening process.

Data relating to standing and sitting postures are presented in Supplement table 5; twenty-two studies reported data about standing posture, five studies about sitting posture, while sixteen studies reported mixed data about both standing and sitting postures.

Study characteristics

The characteristics of the included studies are presented in Supplementary table S3. Twenty-nine studies focused on the intra- and inter-rater reliability of an assessment instrument, and six on test-retest reliability. Eleven studies assessed clinimetric aspects concerning the reliability and validity of one assessment instrument. Seven of the 13 studies about validity concerned concurrent or criterion validity, an item not included in the COSMIN checklist [41]. The eight studies comparing two instruments – neither of which was considered the gold standard – were evaluated using the COSMIN Box (Box H) for criterion validity. One of these two instruments was chosen as the reference standard and considered to be the gold standard, though

with the qualification of 'not a good gold standard'.
Studies concerning aspects of validity were fragmentary, and the diversity of the validity items was large. The study settings were mainly work, laboratory, or school. The total study sample consisted of over 2,600 participants, men and women, aged 5-86 years, with a majority of the sample aged between 18 and 40 years. The study sample of 27 studies was the adult working and/or general population, while six studies focused on children (< 18 years) and only two on musicians.

Study quality
The results of our critical appraisal of the study quality are presented in Supplementary table S4 and figure 2 (QUADAS-II) [40]. The methodological quality of the studies varied considerably (figure 2): 11 out of 41 papers (26.8%) had a score of excellent, with a low risk of bias, low concern regarding applicability, or a maximum of one 'unclear' rating for all items scored. Sixteen studies (39.0%) had a score of moderate, with a maximum of three 'unclear' ratings, one high risk of bias, or major concern regarding applicability. Thus, 66% of the studies had at least a moderate level of methodological quality. The rest of the papers (34%) had a poor level, with at least two high risks of bias or major concerns about applicability and/or at least four 'unclear' ratings.

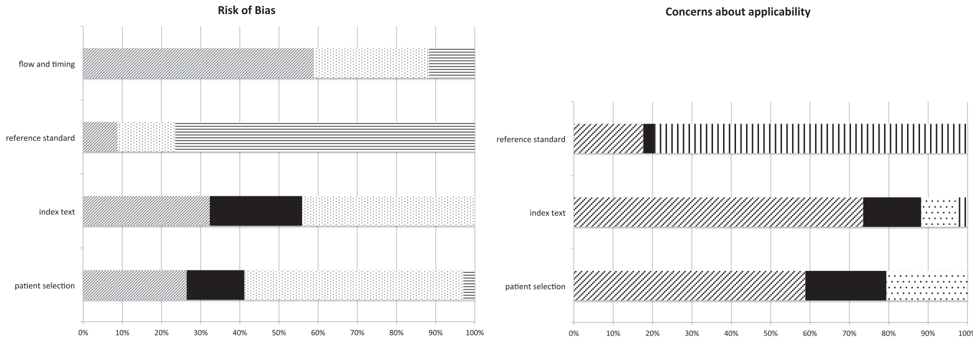

Figure 2: Graphical display of the QUADAS-II assessments (data from Supplement S4).
Legend Risk of bias/applicability: diagonal shading = Low; dotted shading = unclear; solid shading= high; horizontal and vertical shading = not applicable.

Characteristics of assessment instruments
A total of 32 assessment instruments were identified (Supplementary table S5). These were categorized into five groups of assessment methods:

1. Direct body measurement
2. Indirect body measurement (via photograph/video still)
3. Direct visual observation
4. Indirect visual observation (via photograph/video still)
5. Digital measurement: software interpretation of digital 2D-3D photographs/ video stream/stills

In 22 (53.7%) of the studies, a continuous scale was predefined for recording the scores obtained. The head and trunk were the most frequently studied body domains, in 38 studies. The upper and lower extremity domains were less often studied, in 31 and 23 studies respectively, and the least studied was the center of mass domain, in five studies. Twenty-six instruments covered the assessment of three or four body domains, while 10 instruments assessed two adjacent body domains.

Clinimetric values of assessment instruments

The clinimetric values of the different assessment instruments for the observation of posture are listed in Supplementary table S6. For none of the assessment instruments/methods were all items of validity and reliability reported. Two studies reported content or construct validity, while none of the studies reported responsiveness.

Most studies concerned the intra- and inter-rater reliability, with 19 and 29 studies, respectively. We therefore decided to use inter-rater reliability to compare the five categories of assessment instruments, to obtain some indication of one of the clinimetric properties of the instruments.

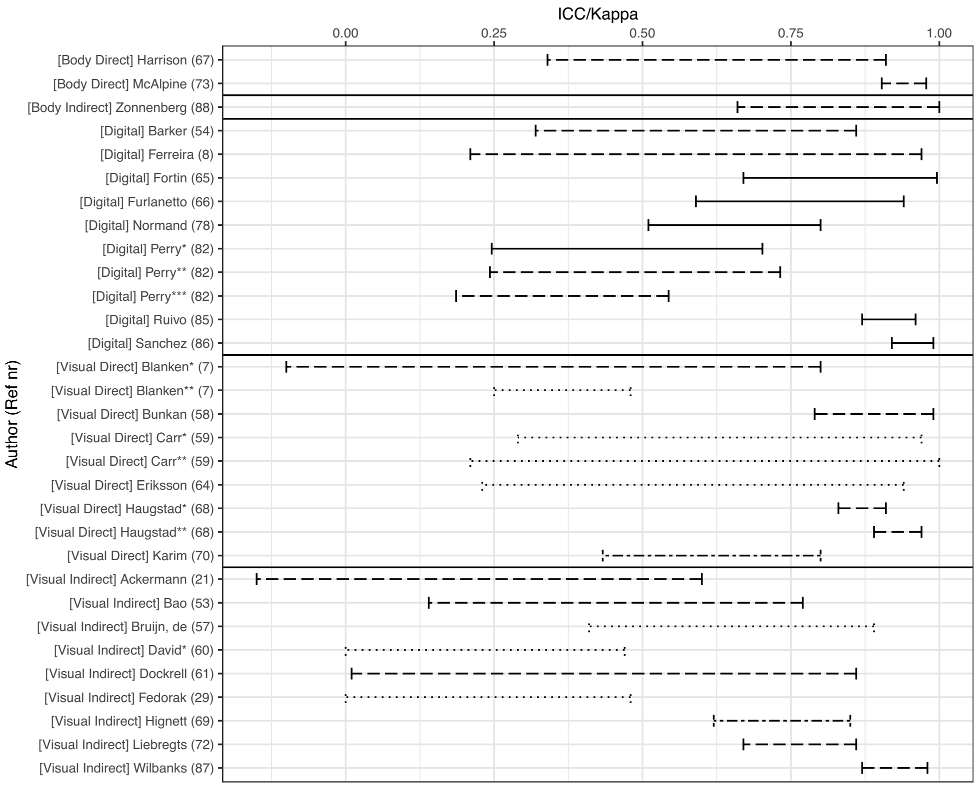

Figure 3: Inter-observer reliability of the assessment instruments per category.
Legend Coefficient type: solid line = ICC; dashed line = ICC (High Quality study); long dashed line = Kappa. Asterisk = multiple instruments in study.

Figure 3 shows a wide dispersion of values in all categories of assessment instruments. The inter-rater reliability values of some posture assessment items are strongly influenced by the nature of these items (e.g. reliability of the assessment of the degree of rotation of a posture domain, or frontal view in comparison to sagittal view). This phenomenon was found for all five categories of observation methods. The 11 studies with the highest methodological quality scores according to QUADAS-II were distributed across two categories of assessment methods: nine studies concerned a digital method and two studies a visual direct method.

Fourteen studies about reliability used measurement error as an indicator; one concerned an indirect body assessment instrument and two concerned indirect visual assessment, while the remaining eleven studies concerned the category of digital assessment instruments. The values for standard error of measurement, minimal detectable difference, or standard deviation were low (0.001–9 mm/0.20–3.8°, with two outliers of 23 mm and 28°). The coefficient of variation varied considerably, with wide confidence intervals [53,62,87].

No relevant differences were observed between the clinimetric and usability values of the measurement methods for either standing or sitting postures.

Clinical usability of the assessment instruments

The clinical usability scores of the different assessment instruments and categories are given in Supplementary table S7 and graphically presented in figure 4. The direct and indirect visual assessment instrument categories had the highest clinical usability scores.

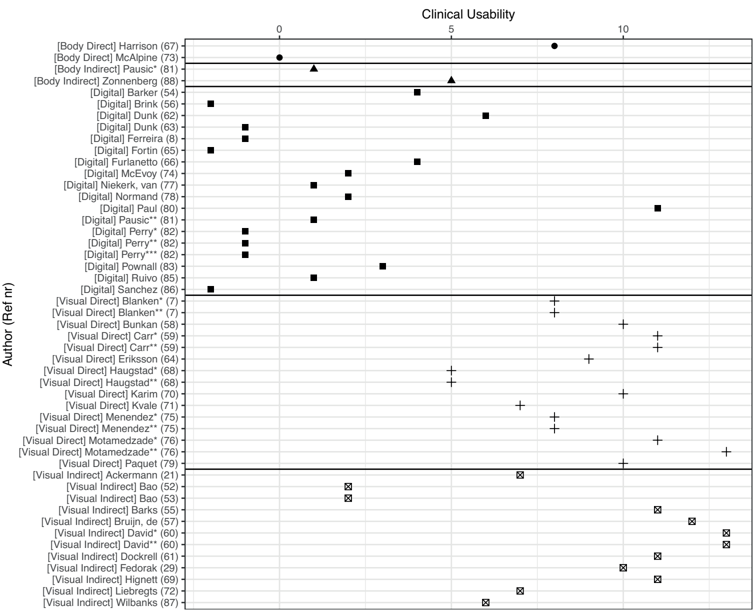

Figure 4: Clinical usability scores of the assessment instruments. Legend Type of instrument: solid circle = body direct; solid triangle = body indirect; solid square = digital; plus sign = visual direct; cross in square sign = visual indirect. Asterisk = multiple instruments in study.

In order to enable a tradeoff between the inter-rater reliability values and the clinical usability of the assessment instrument groups, the reliability data are given for each category in figure 5.

The ideal assessment instrument should have high clinimetric values as well as good clinical usability. The methods that came closest to this ideal were one method in the direct [58] and one in the indirect [69] visual assessment categories. Next to these two studies, one other direct visual method [70] and three visual indirect assessment methods [57,72,87] were identified with a second-best combination score of reliability and usability. Five out of these six methods apply a scaling option for recording the visual assessment. The categories of digital and body assessment instruments scored high on inter-observer reliability but lower on clinical usability.

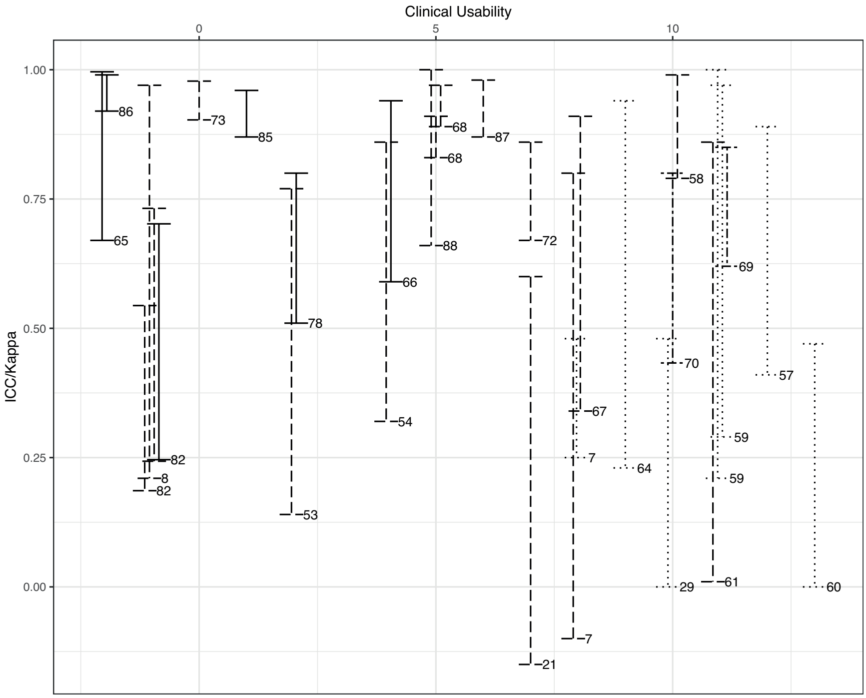

Figure 5: Inter-rater reliability versus clinical usability of the assessment instruments
Legend Coefficient type: solid line = ICC; dashed line = ICC (High Quality study); long dashed line = Kappa.

Aggregation of results

Table 2 summarizes the findings. The number of studies in each of the five categories was small to moderate (range 2-16). Of the five categories, the largest number of studies concerned digital methods and direct and indirect visual methods. All included studies had a cross-sectional design, so ratings of the level of evidence of all methods were restricted to a maximum of 4 points. A wide range of values were found regarding the evaluation of risk of bias, concerns about applicability, consistency of outcomes, and usability scores. The risk of publication bias is presented in Supplementary table S8. We assume that there is a low risk of publication bias because no relevant differences were found between the numbers of smaller and larger studies (≥30 participants) in the different classes of high, moderate, or low reliability values.

Table 2: Summary of findings and recommendations (GRADE)

Studies	Methodolog-ical concerns	Inconsistency of outcomes[2]	Reliability[3]			Validity[5]		Risk of publica-tion Bias[6]	Usability[7] (mean)	Level of recommen-dation[8]
(Ref. no.)	Risk of bias & Applicability concerns[1]		Intra-rater	Inter-rater	Measure ment Error[4]	Criterion & Concurrent V.	Other			
Body measurement direct (n=2)										
Mc Alpine (73)	-2	0	1	1	-	-	-	-1	-1	
Harrison (67)	-1	-1	-	0	-	-	-	0	0	
	-2	**-1**		**1**		**1**		**-1**	**-1**	**D**
Body measurement indirect (n=2)										
Pausic (81)	-2	-	1	-	0	1	-	1	-1	
Zonnenberg (88)	-1	-1	1	0	-	-	-	-1	0	
	-2	**-1**		**0**		**-**		**0**	**-1**	**D**
Visual measurement direct (n=10)										
Blanken (6)	-1	-	-1	-1	-	-	-	-1	0	
Bunkan (57)	-2	0	-	1	-	-	internal consistency: 1	-1	1	
Carr (58)	0	-	-	0-1?	-	-	-	0	1	
Eriksson (64)	-1	-1	-	0	-	-	-	-1	1	
Haugstad (68)	-2	0	-	1	-	-	-	0	0	
Karim (70)	0	-	-	0	-	-	-	0	1	
Kvale (71)	0	-	-	-	-	1	internal consistency: 0	1	0	
Motamedzade (76)	-2	-	-	-	-	0	-	0	1	
Menendez (75)	-1	-	-	-	-	-	-	0	0	
Paquet (79)	-2	-	-	-	-	-	0	-1	0	
	-2	**0**		**0**		**-**		**0**	**1**	**B**
Visual measurement indirect (n=12)										
Ackermann (21)	-1	-1	-1	-1	-	-	-	0	0	
Bao (52)	-1	-	-	-	-	-1	-	1	-1	
Bao (53)	-2	-1	-	-1	-1⁰.	-	-	-1	-1	
Barks (55)	-1	-	-	- (1)#	-	-	content/ construct	-1/0	1	
David (60)	-2	-1	-1	-1-0?	-	1	0	1	1	
de Bruijn (57)	-1	-	1	1	-	-	-	0	1	
Dockrell (61)	-1	-1	0	0	-	-	-	-1	1	
Fedorak (39)	-2	-1	-1	-1	-	-	-	0	1	
Hignett (69)	-2	-	-	0	-	-	-	1	1	
Liebregts (72)	-1	0	0	0	-	1	-	0	0	
Rodby-Bousquet (84)	-1	1	-	1	--1	-	construct validity:1	-1	0	
Wilbanks (87)	-1	1	-	1	-	0	-	-1	0	
	-2	**-1**		**0**		**0**		**0**	**0**	**D**
Digital measurement (n=16)										
Barker (54)	-1	-1	0	0	-	-	-	-1	0	
Brink (55)	-2	-	-	-	-	1	-	-1	-1	
Dunk (62)	-2	-	-	-	-1	-	difference: 1	-1	0	
Dunk (63)	-1	-	-	-	-	-	-	-1	-1	
Ferreira (7)	-2	-1	0	0	-	-	-	-1	-1	
Fortin (65)	0	-1	1	1	0.	-	-	0	-1	
Furlanetto (66)	0	0	0	0	0	-	-	0	-	
Mc Evoy (74)	0	-	-	-	-	-	-	0	-1	
Niekerk van (77)	-1	-	-	-	-1	-	-	0	-1	
Normand (78)	0	-1	0	0	0	-	-	0	-1	
Paul (80)	0	-	-	-	-	-	difference: 1	-1	1	
Pausic (81)	-2	-	1	-	0	-	construct validity: 1	1	-1	
Perry (82)	0	-1	1	-1-0?	-1.	-	-	-1	-1	
Pownall (77)	-2	-1	-	0	0	-	-	-1	2	
Ruivo (85)	0	1	-	1	0	-	-	1	-1	
Sanchez (86)	0	1	-	1	0	-	-	-1	-1	
	-1	**-1**		**0**		**1**		**-1**	**-1**	**D**

[1] Interpretation of risk of bias and concerns about applicability (score on QUADAS II): 0 = ≤ 1 uncertain score/study; -1 = ≤3 uncertain scores or ≤ 1 high score/study; -2 = > 4 uncertain scores or > 1 high score/study; at group level, the mean value is used for deducting or adding points for this category; the values are rounded.

[2] Inconsistency of outcomes (inter-rater reliability): 1 = ≥ 75% of correlations coefficients ≥ 0.75/study; 0 = ≥ 75% of correlations coefficients 0.5 ≤ correlations coefficients < 0.75/study; -1 = lower percentages of correlation coefficients or lower inter-rater reliability values.

[3] Range of scores: 1 = ≥ 75% of number of values ≥ 0.75; 0 = 75% of values 0.5 ≤ values < 0.75; -1 = ≥ 75% of values < 0.5; at group level, the mean value is used for deducting or adding points for this category, the values are rounded. # = Test-retest value in the absence of inter-rater reliability value.

[4] Measurement errors: 0= ≤ 10 mm or ≤ 10 degrees; -1= > 10 mm or > 10 degrees.

[5] Criterion/Concurrent/Construct validity/Internal consistency: 1= correlation value ≥| ±0.7| ; 0 = | ±0.3| ≤ correlation value ≤| ±0.7| ; -1= correlation value ≤| ±0.3| ; 1 = Differences ≤ 5^0 or ≤ 10 mm; 0 = Differences > 5^0 or >10 mm.

[6] Risk of publication bias: 1= > 100 participants/study; 0 = 30-100 participants/study; -1= < 30 participants/study.

[7] Usability value (Supplement table S7): 1 = ≥ 9 points/study; 0 = 3 ≤ points/study < 9; -1 = < 3 points/study; at group level, the mean value is used for deducting or adding points for this category; the values are rounded.

[8] Level of recommendation: The mean sum score (range 1-4) of the initial score at group level and the deduced and added points at group level.

Discussion

The aim of this systematic review was to provide an overview of the clinimetric properties of assessment methods for static standing and/or sitting posture of musicians in routine clinical settings, and to interpret the findings for clinical practice. Thirty-two instruments for the clinical assessment of sitting and/or standing posture were identified. The instruments were divided over five categories: assessment methods using direct body measurements, indirect body measurements (via photographs/video stills), direct visual observations, indirect visual observations, and digital assessment methods (using any form of software to collect information from photographs/video stills).

The following five tentative conclusions were drawn. Firstly, the direct and indirect visual assessment instruments, using a rating scale to record the aspects of posture, seem to have the best combination of inter-rater reliability and usability. Secondly, little and fragmentary data was found about validity-related aspects, and no data about responsiveness. Thirdly, the inter-rater reliability values of some posture assessment items are strongly influenced by the nature of these items (e.g. the reliability of assessing the degree of rotation of a posture domain, or frontal view in comparison to sagittal view). This phenomenon applied to all five categories of observation methods. Fourthly, the measurement error values (standard error of measurement, minimal detectable difference, and/or standard deviation) are generally low (<40 or <9 mm). Fifthly, a weak recommendation (GRADE level B), based on a moderate level of evidence, can be made for clinicians to use the direct or indirect visual observation method, using a rating scale recording form, and having it administered by a trained observer. The conclusions are partly in line and sometimes conflicting with conclusions from other reviews. [8,11,22,30-39] The most similar systematic review [22] about the assessment of biomechanical exposures at work (evaluating both global posture and individual body domains) concluded that none of the observation methods is superior to the others, and that global body postures are the most reliable to measure. Our present review comes to different conclusions, , as shown by the wide range of values in inter-reliability values over the five categories. However, the reliability values for the direct and indirect body measurement methods are higher than those for the other categories, and there are clear differences between the five categories in the trade-off between inter-rater reliability and usability values (with the highest values for both being found for the direct and indirect visual method). On the other hand, the wide range of values we found means that, like Takala et al. [22], we can only draw tentative conclusions.

There are several reasons why we need to be careful in drawing firm conclusions. The main reason is that for most assessment methods, it is only the reliability which has been thoroughly studied, while the other clinimetric properties are unknown. Another reason is that it is not known to what extent the clinimetric properties of the instruments designed to assess the posture of one body domain are comparable with those of instruments designed for the assessment of all of the major body domains or the global body posture. Our full-text screening excluded 200 papers about assessing the posture of a single body domain. The systematic review by Takala et al. [22] included both single body domain and global posture observation studies, but it is

not clear how many articles were identified in their study (in the absence of a flow diagram of the screening procedure). Hence, comparing the outcomes of our study with those of the study by Takala et al. [22] is difficult.
Fortin et al. [11] concluded in their narrative review that the quantitative assessment of global posture is performed most accurately and rapidly by measuring body angles from photographs. This conclusion might be based on their studies of single body domains. The results of our review (about global posture) are not in agreement with this. We assume that the digital assessment methods will be more suitable to assess global posture in a reliable and usable way than the visual assessment methods, especially with the advent in recent years of a wide range of new posture assessment apps and photogrammetry software (sometimes freely available on the internet) [36,37]. These are promising assessment methods which might be expected to yield high clinimetric and clinical usability values in the near future. These new methods have, however, not yet been tested in validation studies. Moreover, the application of photogrammetry in postural evaluation is directly dependent on both the collection procedures and the mathematical methods used to provide measurements. In line with Fortin et al. [11] and Furlanetto et al. [33], we found that the postural evaluation software that is used varies greatly among the studies, with often no explanation about the methods used to generate the results. This makes it difficult to interpret data synthesis rules within these 'black boxes', and the software is often not accessible [33].

In the studies included in our review, the measurement error values at the participant group level per study were low, but the confidence intervals were wide. This is due to a combination of variations in marker placement, differences in parameter definitions, body position, perspective error (due to camera position), and especially biological variability (particularly among children due to anthropometric and motor control immaturity). No conclusion can therefore be drawn about the ecological validity or the interpretation of these values for individuals in a clinical setting,

The major strengths of our study are the large number of screened and included records, the fact that our conclusions are based on studies of which the majority had moderate to good methodological quality, and the fact that all procedures as much as possible followed the international standards for performing (Cochrane, PRISMA) [27,50] and reporting (AMSTAR) [51] systematic reviews, and the systematic and explicit approach (GRADE) [44-49] we used to judge the quality of evidence and strength of recommendations for clinicians.

Potential limitations of the study are, in theory, a risk of selection bias of articles and a possible bias in the process of interpreting and aggregating the findings. The risk of selection bias due to publication bias can be assumed to be low, because no relevant differences were found between the numbers of small and large studies assigned to the various classes of correlation coefficients for the outcomes [43]. There were many choices to be made during the process of interpretation and aggregation, and each of these choices required weighing the evidence. The guidelines offer no solution to this problem. An explicit description of the arguments for our choices is provided in this article as much as possible.

Another limitation of the study is that the usability values of the assessment instruments were based on a self-constructed scoring list. We are aware of the subjective nature of this list. As far as we know, there is no objective way to score clinical usability. Before the start of the study we asked several clinicians to review the scoring list. Most of the discussions and subjective views concerned the item of 'cost price of the assessment instrument'. There were different opinions about the criteria for the various intervals, e.g. depending on the clinical setting (in that the estimation and acceptability of the costs for an instrument used by many therapists in rehabilitation centers differs from those of the same instrument used by a single therapist in a peripheral physical therapy practice). The cost-price item was also difficult to divide into classes, as little information about it was presented in the papers we reviewed. The bias due to this uncertainty might be that the actual clinical usability scores may be one or two points higher, especially in the indirect assessment categories and the digital category (as these methods use technical support in some form).

Presenting separate data for the sitting or standing posture was only partly possible in our review. Several studies presented combined data for both postures, while in others it was unclear whether either postures or just one had been included. The consequences of this omission are small, however, as the clinimetric and usability values of the assessment methods are similar for both postures.

It is not clear to what extent the conclusions of our review are generalizable to subpopulations; we found insufficient papers about e.g. age groups, patients versus healthy people. What little information we could retrieve from the studies does not appear to show relevant differences, except that a lower level of reliability has to be taken into account with younger children, as their balance maintenance is less mature than that in adults. [74]

In line with Takala et al. [22] and Furlanetto et al. [33], we recommend that selecting a clinical assessment method for posture should be based on the clinician's purpose. Based on our review it seems best to opt for the direct or indirect visual assessment methods, as these provide the best combination of clinimetric and usability values. However, these instruments are less appropriate for the quantification and evaluation of posture, and are less responsive to change.

The direct or indirect visual observation method is best for situations where a (quick) qualitative observation of posture is required, and/or an estimated quantitative and/or qualitative evaluation of posture (e.g. classification in a rating scale with 3 classes).

In view of the near absence of studies evaluating the construct validity or predictive validity of assessment instruments for static sitting and standing posture, recommendations to clinicians should be made with caution if an assessment method is to be used to detect postures carrying a risk of musculoskeletal complaints. Little information was also found about the criterion validity aspects. In other words, assumptions about what is relevant in assessing and judging static postures in terms of carrying a risk of musculoskeletal complaints should be critically reconsidered. In terms of the GRADE framework, the level of recommendation for most diagnostic instruments is often low, because data about these aspects are scarce [48]. Based on these arguments we tentatively recommend the use of the standardized direct or indirect visual observation method for the assessment of static posture. The use of other methods in clinical practice is not supported by the results of our review.

Conclusion

Based on a moderate level of evidence according to the GRADE framework, a weak recommendation can be made for using the direct or indirect visual assessment method (with posture recorded as interval scores by a trained observer) to assess sitting and/or standing posture in daily clinical practice. Little and fragmentary information was found about validity-related aspects, and no data about responsiveness. For all five categories of observation methods 'of musicians, the inter-rater reliability values of some posture assessment items are strongly influenced by the nature of these items.

Acknowledgments

We would like to thank Prof. Henrica CW de Vet, PhD and Ms. Wieneke Mokkink, PhD for their theoretical support in the construction phase of the study, and Mr. Guus van den Brekel & Mrs. Sjoukje van der Werf for their technical advice and support in the search strategy phase of the study.

Reference list

[1] Kok Lm, Huisstede BM, Voorn VM, Schoones JW, Nelissen RG. The occurence of Musculoskeletal complaints among professional musicians: a systematic review. Int Arch Occup Environ Health. 2016;89(3):373-396.

[2]. Cailliet R. Abnormalities of the sitting postures of musicians. Med Probl Perform Art.1990;5:131-135.

[3]. National Research Council and the Institute of Medicine. Musculoskeletal disorders and the workplace: low back and upper extremities. Panel on Musculoskeletal Disorders and the Workplace. Commission on behavioral and Social Sciences and Education. Washington, DC: National Academy Press. 2001.

[4]. Pascarelli EF, Hsu YP. Understanding work-related upper extremity disorders: Clinical findings in 485 computer users, musicians, and others. J Occup Rehabil. 2001;11:1-21.

[5]. Picavet HS, Schouten JS. Musculoskeletal pain in the Netherlands: Prevalences, consequences and risk groups, the DMC(3)-study. Pain. 2003;102:167-178.

[6]. Windt van der DA, Thomas E, Pope DP, et al. Occupational risk factors for shoulder pain: A systematic review. Occup Environ Med. 2000;57:433-442.

[7]. Blanken WCG, Rijst van der H, Mulder PGH, Eijsden van - Besselink MDF, Lankhorst GJ. Interobserver and intraobserver reliability of postural examination. Med Probl Perform Art. 1991;6:93-97.

[8]. Ferreira EAG, Duarte M, Maldouado EP, Burke TN, Marques AP. Postural assessment software (PAS/SAPO): validation and reliability. Clinics 2010;65:675-681.

[9]. Abréu-Ramos AM, Brandfonbrener AG. Lifetime prevalence of upper-body musculoskeletal problems in a professional-level symphony orchestra: Age, gender, and instrument-specific results. Med.Probl.Perform.Art. 2007;22:97-104.

[10]. Brandfonbrener A,G. History of playing-related pain in 330 university freshman music students. Med Probl Perform Art. 2009;24:30-36.

[11]. Fortin C, Ehrmann-Feldman D, Cheriet F, Labelle H. Clinical methods for quantifying body segment posture: A literature review. Dis and Rehab. 2011;33:367-383.

[12]. Woldendorp KH, Tijsma A, Boonstra AM, Arendzen JH, Reneman MF. No association between posture and musculoskeletal complaints in a professional bassist sample. EJP.2016;20:399-407.

[13]. Aghilinejad M, Azar NS, Ghasemi MS, Dehghan N, Mokamelkhah EK. An ergonomic intervention to reduce musculoskeletal discomfort among semiconductor assembly workers. Work. 2016;54:445-450.

[14]. Kilroy N, Dockrell S. Ergonomic intervention: its effect on working posture and musculoskeletal symptoms in female biomedical scientists. Br J Biomed Sci. 2000;57:199-206.

[15]. Pillastrini P, Mugnai R, Bertozzi L, et al. Effectiveness of an ergonomic intervention on work-related posture and low back pain in video display terminal operators: a 3 year cross-over trial. Appl Ergon. 2010;4:436-443.

[16]. Rosario JL. Relief from back pain through postural adjustment: a controlled clinical trial of the immediate effects of muscular chains therapy (MCT). Int J Ther Massage Bodywork. 2014;7:2-6.

[17]. Goldberg CJ, Kaliszer M, Moore DP, Fogarty EE, Dowling FE. Surface topography, cobb angles, and cosmetic change in scoliosis. Spine. 2001;26:E55-E63.

[18]. Nault ML, Allard P, Hinse S, et al. Relations between standing stability and body posture parameters in adolescent idiopathic scoliosis. Spine. 2002;27:1911-1917.

[19]. Pazos V, Cheriet F, Song L, Labelle H, Dansereau J. Accuracy assessment of human trunk surface 3D reconstructions from an optical digitising system. Med Biol Eng Comput. 2005;43:11-15.

[20]. Zabjek KF, Leroux MA, Rivard CH, Prince F. Evaluation of segmental postural characteristics during quiet standing in control and idiopatic scoliosis patients. Clin Biomech. 2005;20:483-490.

[21]. Ackermann BJ, Adams R. Interobserver reliability of general practice physiotherapists in rating aspects of the movement patterns of skilled violinists. Med Probl Perform Art. 2004;19:3-11.

[22]. Takala EP, Pehkonen I, Forsman M, Hansson GA, Mathiassen SE, Neumann WP, et al. Systematic evaluation of observational methods assessing biomechanical exposures at work. Scan J Work, Environ & Health. 2010;36(1):3-24.

[23]. Raine S, Twomey L. Attributes and qualities of human posture and their relationship to dysfunction or musculoskeletal pain. Crit Rev Phys and Rehabil Med. 1994;6:409-437.

[24]. Yates T, Wilmot EG, Davies MJ, et al. Sedentary behavior: what's in a definition? Am.J.Prev.Med. 2011;40:e33-e34.

[25]. Dictionary. (n.d.). Retrieved August 18, 2016 from https://www.merriam-webster.com/dictionary/dictionary/sit.

[26]. Mens en Werk, Arbo catalogus. Retrieved August 18, 2016 from http://www.ii-mensen werk.nl/Upload/Arbocatalogus.

[27]. The Cochrane Collaboration. Handbook for DTA reviews. Retrieved November 12, 2014 from http://srdta.cochrane.org/handbook -dta-reviews.

[28]. Vet de HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine. 3rd printing. Cambridge: Cambridge University Press; 2014.

[29]. Fedorak C, Ashworth TS, Marshall J, Paull H. Reliability of the visual assessment of cervical and lumbar lordosis: How good are we? Spine. 2003;28:1857-1859.

[30]. Baxter ML, Milosavljevic S, Ribeiro DC, Hendrick P, McBride D. Psychometric properties of visually based clinical screening tests for risk of overuse injury. Phys Ther Rev. 2014;19:213-219.

[31]. Burt S, Punnett L. Evaluation of interrater reliability for posture observations in a field study. Appl Erg. 1999;30:121-135.

[32]. Field D, Livingstone R. Clinical tools that measure sitting posture, seated postural control or functional abilities in children with motor impairments: a systematic review. Clin Rehab. 2013; 27(11):994-1004.

[33]. Furlanetto TS, Sedrez JA, Tarragô Condotti C, Fagundes Loss J. Photogrammetry as a tool for postural evaluation of the spine: A systematic review. World J Orthop. 2016;7(2):136-148.

[34]. Gallagher S. Physical limitations and musculoskeletal complaints associated with work in unusual or restricted postures: a literature review. J Safety Research. 2005;36:51-61.

[35]. Kilbom A. Assessment of physical exposure in relation to work-related musculoskeletal disorders: what information can be obtained from systematic observations? Scan J Work Environ Health. 1994;20 Special issue:30-45.

[36]. Krawczky B, Pacheco AG, Mainenti MRM. A systematic review of the angular values obtained by computerized photogrammetry in sagittal plane: a proposal for reference values. J Manipulative Physiol Ther. 2014;37:269-275.

[37]. Linn JM. Using digital image processing for the assessment of postural changes and movement patterns in bodywork clients. J Bodywork Mov Ther. 2001;5:11-20.

[38]. Singla D, Veqar Z. Methods of postural assessment for sports persons. J Clin Diagn Research. 2014;8:LE01-LE04.

[39]. Wong WY, Wong MS, Lo KH. Clinical applications of sensors for human posture and movement analysis: a review. Prosthet Orthot Int. 2007;31:62-75.

[40]. Whiting PF, Rutjes AWS, Westwood ME, et al. A revised instrument for the Quality Assessment of Diagnostic Accuracy Studies. Ann Intern Med. 2011;155:529-536.

[41]. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010;63:737-745.

[42]. Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LM, Vet de HCW. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res. 2012;21:651-657.

[43]. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. 2nd ed. Upper Saddle River: J. Alexander; 2000.

[44]. Atkins D, Eccles M, Flottorp S, et al. Systems for grading the quality of evidence and the strength of recommendations I: critical appraisal of existing approaches. The GRADE Working Group. BMC Health Serv Res. 2004;4:38. doi 10.1186/1472-6963-4-38

[45]. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336:924-926.

[46]. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: what is "quality of evidence" and why is it important to clinicians. BMJ. 2008;336:995-998.

[47]. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: going from evidence to recommendations. BMJ. 2008;336: 1049-1051.

[48]. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: grading quality of evidence and strength of recommendations for diagnostic tests and strategies. BMJ. 2008;336:1106-1110.

[49]. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: incorporating considerations of resources use into grading recommendations. BMJ. 2008;336:1170-1173.

[50]. Moher D, Librati A, Tetzlaff J, Altman DG, and the PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med. 2009;151:264-269.

[51]. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement instrument to access the methodological quality of systematic reviews. BMC Med Research Methodol. 2007;7:10. doi 10.1186/1471-2288-7-10.

[52]. Bao S, Howard N, Spielholz P, Silverstein B. Two posture analysis approaches and their application in a modified rapid upper limb assessment evaluation. Ergonomics. 2007;50:2118-2136.

[53]. Bao S, Howard N, Spielholz P, Silverstein B.Polissar N. Interrater reliability of posture observations. Hum Factors. 2009;51:292-309.

[54]. Barker SP, Tremback-Ball A, Sterner K. Reliability of posture measurements taken with photo analysis. J Women Health Physical Ther. 2006;30:2.

[55]. Barks L, Luther SL, Brown LM, Schulz B, Bowen ME, Powell-Cope G. Development and initial validation of the Seated Posture Scale. JRRD. 2015;52(2):201-210.

[56]. Brink Y, Louw Q, Grimmer K, Schreve K, Westhuizen van der G, Jordaan E. Development of a cost effective three-dimensional posture analysis tool: validity and reliability. BMC Musculoskelet Disord. 2013;14:335.

[57]. Bruijn de I, Engels JA, van der Gulden JW. A simple method to evaluate the reliability of OWAS observations. Appl Ergon. 1998;29:281-283.

[58]. Bunkan BH, Moen O, Opjordsmoen S, Ljunggren AE, Friis S. Interrater reliability of the comprehensive body examination. Physioth Theory Pract. 2002;18:121-129.

[59]. Carr EK, Kenney FD, Wilson-Barnett J, Newham DJ. Inter-rater reliability of postural observation after stroke. Clin Rehab. 1999;13:229-242.

[60]. David G, Woods V, Li G, Buckle P. The development of the Quick Exposure Check (QEC) for assessing exposure to risk factors for work-related musculoskeletal disorders. Appl Ergon. 2008;39:57-69.

[61]. Dockrell S, O'Grady E, Bennett K, et al. An investigation of the reliability of Rapid Upper Limb Assessment (RULA) as a method of assessment of children's computing posture. Appl Ergon.2012;43:632-636.

[62]. Dunk NM, Chung YY, Compton DS, Callaghan JP. The reliability of quantifying upright standing postures as a baseline diagnostic clinical tool. J Manipulative Physiol Ther. 2004;27:91-96.

[63]. Dunk NM, Lalonde J, Callaghan JP. Implications for the use of postural analysis as a clinical diagnostic tool: reliability of quantifying upright standing spinal postures from photographic images. J Manipulative Physiol Ther. 2005;28:386-392.

[64]. Eriksson EM, Mokhtari M, Pourmotamed L, Holmdahl L, Eriksson H. Inter-rater reliability in a resource-oriented physiotherapeutic examination. Physioth Theory Pract. 2000;16:95-103.

[65]. Fortin C, Feldman DE, Cheriet F, Gravel D, Gauthier F, Labelle H. Reliability of a quantitative clinical posture assessment tool among persons with idiopathic scoliosis. Physiother.2012;98:64-75.

[66]. Furlanetto T S, Candotti CT, Sedrez JA, Noll M, Loss JF. Evaluation of the precision and accuracy of the DIPA software postural assessment protocol. Eur J Physiother. 2017;19(4):179-184.

[67]. Harrison AL, Barry-Greb T, Wojtowicz G. Clinical measurement of head and shoulder posture variables. J Orthop Sports Phys Ther. 1996;23:353-361.

[68]. Haugstad GK, Haugstad TS, Kirste U et al. Reliability and validity of a standardized Mensendieck physiotherapy test (SMT). Physiother Theory Pract. 2006;22:189-205.

[69]. Hignett S, McAtamney L. Rapid entire body assessment (REBA). Appl Ergon. 2000;31:201-205.

[70]. Karim A, Millet V, Massie K, Olson S, Morganthaler A. Inter-rater reliability of a musculoskeletal screen as administered to female professional contemporary dancers. Work. 2011;40:281-288.

[71]. Kvale A, Bunkan BH, Opjordsmoen S, Ljunggren AE, Friis S. Development of the Posture domain in the Global Body Examination (GBE). Adv Physiother. 2010;12:157–165.

[72]. Liebregts J, Sonne M, Potvin JR. Photograph-based ergonomic evaluations using the Rapid Office Strain Assessment (ROSA). Appl Ergon. 2016;52:317-324.

[73]. McAlpine RT, Betaany-Saltikov JA, Warren JG. Computerized back postural assessment in physiotherapy practice: Intra-rater and inter-rater reliability of the MIDAS system. J Back Musculoskelet Rehabil. 2009;22:173-178.

[74]. McEvoy MP, Grimmer K. Reliability of upright posture measurements in primary school children. BMC Musculoskelet Disord. 2005;6:35. doi: 10.1186/1471-2474-6-35

[75]. Menendez CC, Amick III BC, Chang CC. et al. Evaluation of two posture survey instruments for assessing computing postures among college students. Work. 2009;34:421-430.

[76]. Motamedzade M, Ashuri MR, Golmohammadi R, Mahjub H. Comparison of ergonomic risk assessment outputs from rapid entire body assessment and quick exposure check in an engine oil company. J Res Health Sci. 2011;11:26-32.

[77]. Niekerk van SM, Louw Q, Vaughan C, Grimmer-Somers K, Schreve K. Photographicmeasurement of upper-body sitting posture of high school students: a reliability and validity study.BMC Musculoskelet Disord. 2008;9:113. doi: 10.1186/1471-2474-9-113.

[78]. Normand MC, Descarreaux M, Harrison DD, et al. Three dimensional evaluation of posture in standing with the PosturePrint: an intra- and inter-examiner reliability study. Chiropr Osteopat.2007;15: 15. doi: 10.1186/1746-1340-15-15

[79]. Paquet VL, Punnett L, Buchholz B. Validity of fixed-interval observations for postural assessment in construction work. Appl Ergon. 2001;32:215-224.

[80]. Paul JA, Douwes M. Two-dimensional photographic posture recording and description: a validity study. Appl Ergon. 1993;24:83-90.

[81]. Pausić J, Pedisić Z, Dizdar D. Reliability of a photographic method for assessing standing posture of elementary school students. J Manipulative Physiol Ther. 2010;33:425-431.

[82]. Perry M, Smith A, Straker L, Coleman J, O'Sullivan P. Reliability of sagittal photographic spinal posture posture assessment in adolescents. Adv Physiother. 2008;10:66-75.

[83]. Pownall PJ, Moran RW, Stewaret Am. Consistency of standing and seated posture of asymptomatic male adults over a one-week interval: a digital camara analysis of multiple landmarks. Int J Osteopath Med. 2008;11:43-51.

[84]. Rodby-Bousquet E, Ágústsson A, Jónsdóttir G, Czuba T, Johansson AC, Hägglund G. Interrater reliability and construct validity of the Posture and Postural Ability Scale in adults with cerebral palsy in supine, prone, sitting and standing positions. Clin Rehab. 2014;28:82-90.

[85]. Ruivo RM, Pezarat-Correia P, Carita AI. Intrarater and interrater reliability of photographic measurement of upper-body standing posture of adolescents. J Manipulative Physiol Ther. 2015;38(1):74-80.

[86]. Sánchez MB), Loram I, Darby J, Holmes P, Butler PB. A video based method to quantify posture of the head and trunk in sitting. Gait Post. 2017;51:181-187.

[87]. Wilbanks S. To determine the reliability of a digital photo-based posture assessment in standing. Arch Phys Med Rehab. 2016;97(12):e6.

[88]. Zonnenberg AJ, Maanen van CJ, Elvers JW, Oostendorp RA. Intra/interrater reliability of measurements on body posture photographs. Cranio.1996;14:326-331.

Supplement table 1

Search string (Web of Science)

(TI=Analyze OR TS=Analyze OR TI=Analyse OR TS=Analyse OR TI=Asses OR TS=Asses OR TI=Assessing OR TS=Assessing OR TI=Diagnose OR TS=Diagnose OR TI=Evaluation OR TS=Evaluation OR TI=Evaluate OR TS=Evaluate OR TI=Examine OR TS=Examine OR TI=Inspect OR TS=Inspect OR TI=Measure OR TS=Measure OR TI=Observe OR TS=Observe OR TI=Screen OR TS=Screen OR TI=Assessment OR TS=Assessment OR TI=Analy* OR TS=Analy* OR TI=Checklist OR TS=Checklist OR TI=Physical examination OR TS=Physical examination OR TI=Examination OR TS=Examination OR TI=Instrument OR TS=Instrument OR TI=Inspection OR TS=Inspection OR TI=Measurement OR TS=Measurement OR TI=Method OR TS=Method OR TI=Observation OR TS=Observation OR TI=Registration OR TS=Registration OR TI=Screening OR TS=Screening OR TI=Standard OR TS=Standard OR TI=System OR TS=System OR TI=Technique OR TS=Technique OR TI=Test OR TS=Test)

AND

(TI=Postu* OR TS=Postu* OR TI=(Human AND body) OR TS=(Human AND body) OR TI=Body-segment OR TS=Body-segment OR TI=(Human AND alignment) OR TS=(Human AND alignment) OR TI=(Joint AND position) OR TS=(Joint AND position))

AND

(TI=Reproducibility OR TS=Reproducibility OR TI=Validity OR TS=Validity OR TI=(Reproducibility AND of AND Results) OR TS=(Reproducibility AND of AND Results) OR TI=Specificity OR TS=Specificity OR TI=Sensitivity OR TS=Sensitivity OR TI=(positive AND predictive AND value) OR TS=(positive AND predictive AND value) OR TI=(Negaitive AND predictive AND value) OR TS=(Negative AND predictive AND value) OR TI=Reliabil* OR TS=Reliabil* OR TI=reliabilty OR TS=reliabilty OR TI=Feasibility OR TS=Feasibility OR TI=Responsitivity OR TS=Responsitivity OR TI=Usability OR TS=Usability OR TI=Practical* OR TS=Practical* OR TI=Clinical* OR TS=Clinical* OR TI=Readability OR TS=Readability OR TI=Readable OR TS=Readable OR TI=Legible OR TS=Legible OR TI=Skillful OR TS=Skillful OR TI=Skilfull OR TS=Skilfull OR TI=(Execution AND time) OR TS=(Execution AND time) OR TI=Consistency OR TS=Consistency OR TI=(Ease AND of AND use) OR TS=(Ease AND of AND use) OR TI=Characteristics OR TS=Characteristics OR TI=Properties OR TS=Properties OR TI=Efficacy OR TS=Efficacy OR TI=Effectiveness OR TS=Effectiveness OR TI=Valid* OR TS=Valid* OR TI=Usage OR TS=Usage OR TI=Repeatibility OR TS=Repeatibility OR TI=Agreement OR TS=Agreement OR TI=inter-rater-reliability OR TS=inter-rater-reliability OR TI=interraterreliability OR TS=interraterreliability OR TI=inter-raterreliability OR TS=inter-raterreliability OR TI=interrater-reliability OR

TS=interrater-reliability OR TI=intra-rater-reliability OR TS=intra-rater-reliability OR TI=intraraterreliability OR TS=intraraterreliability OR TI=intra-raterreliability OR TS=intra-raterreliability OR TI=intrarater-reliability OR TS=intrarater-reliability)

NOT

TI=(pharmaceutical OR preparations OR drugs OR drugs0 OR drugs1 OR drugs2 OR drugs2004 OR drugs2010 OR drugs99 OR drugsafdeling OR drugsafety OR drugsagainststaphylococcus OR drugsand OR drugsandhomeless OR drugsare OR drugsatfda OR drugsbiomarkersbasic OR drugsc OR drugscan OR drugscapes OR drugscene OR drugscience OR drugscontroller OR drugscope OR drugscreen OR drugscreening OR drugscreens OR drugsd OR drugsdepartment OR drugsduring OR drugseller OR drugsellers OR drugselling OR drugsense OR drugsensitive OR drugsf OR drugsgebruik OR drugsgebruikers OR drugsgerelateerde OR drugshas OR drugshin OR drugshop OR drugshulpverlening OR drugsim OR drugsin OR drugsincluding OR drugsindermatology OR drugsite OR drugsmet OR drugsmeta OR drugsnon OR drugson OR drugsopioid OR drugsoverdosering OR drugspace OR drugsprotection OR drugss OR drugsspuitster OR drugstar OR drugstat OR drugster OR drugsthat OR drugsthere OR drugsthree OR drugsthrough OR pharmacies OR pharmacie OR drugstore OR drugstores OR drugsurv OR drugsusceptibility OR drugsusceptible OR drugsverslaafden OR drugsverslaving OR drugswere OR drugswhich OR drugsyy2013 OR medicine OR transfusion OR veterinary OR molecular OR travel OR infectious OR emergency OR biochemical OR biochemicalcell OR biochemicalcharacterization OR biochemicale OR biochemicalengineering OR biochemicalical OR biochemicall OR biochemically OR biochemicallyweighted OR biochemicalneural OR biochemicalparameters OR biochemicalphysiological OR biochemicals OR biochemicalstudies OR biochemicaly OR lung OR lungs OR pulmonary OR pulmonaryand OR pulmonaryartery OR pulmonarycomplications OR pulmonaryconsultants OR pulmonarycosis OR pulmonarydisease OR pulmonaryedema OR pulmonaryeembolism OR pulmonaryfibrosis OR pulmonaryfibrosisresearch OR pulmonarygas OR pulmonaryhypertension OR pulmonaryinfection OR pulmonarymdct OR pulmonarymedicine OR pulmonarymetastases OR pulmonarynodules OR pulmonaryplegia OR pulmonaryplegic OR pulmonaryresection OR pulmonarytelerehabilitation OR pulmonarytelerehabilitaton OR pulmonarytuberculosis OR pulmonaryuniversity OR pulmonaryvascular OR pulmonaryvein OR pulmonaryvenous OR disease OR invasive OR invasiveacacia OR invasiveapproach OR invasivebladder OR invasivecardiac OR invasivecollection OR invasived OR invasivedoc OR invasiveive OR invasiveless OR invasivelesslessbenefit OR invasivelly OR invasively OR invasivem OR invasiven OR invasivenature OR invasivencss OR invasivene OR invasivenes OR invasiveness OR invasivenss OR invasiveo OR invasivep OR invasivepossible OR invasiver OR invasives OR invasivesess OR invasivesness OR invasivesurgery OR invasivetka OR invasivetype

OR medicine1 OR medicine10290 OR medicine10department OR medicine11 OR medicine11department OR medicine1550 OR medicine16 OR medicine1600 OR medicine1971 OR medicine2 OR medicine2041 OR medicine209 OR medicine21 OR medicine295 OR medicine2department OR medicine2lsh OR medicine2music OR medicine2philip OR medicine2school OR medicine3 OR medicine310 OR medicine333 OR medicine34 OR medicine354 OR medicine3800 OR medicine3associate OR medicine3duke OR medicine4 OR medicine4department OR medicine4division OR medicine4robert OR medicine5center OR medicine5section OR medicine6department OR medicinea OR medicineaalborg OR medicineaarhus OR medicineac OR medicineacademic OR medicineafrica OR medicineajou OR medicineakershus OR medicineal OR medicinealbuquerque OR medicineand OR medicineankara OR medicinearistotle OR medicineat OR medicineatlanta OR medicineavepark OR medicineb OR medicinebaltimore OR medicinebarts OR medicinebaystate OR medicinebeth OR medicinebiological OR medicinebiostatisticspathologyendocrine OR medicinebispebjerg OR medicinebrigham OR medicinebronx OR medicinec OR medicinecanterbury OR medicinecardiff OR medicinecardiothoracic OR medicinecase OR medicinecenter OR medicinecentre OR medicinecheng OR medicinechiba OR medicinechicago OR medicinechief OR medicinechinese OR medicinechung OR medicineclayton OR medicineclinical OR medicineco OR medicinecollege OR medicinecolorado OR medicineconway OR medicinecork OR medicinecould OR medicined OR medicinedanish OR medicinedavid OR medicinedentistry OR medicinedenver OR medicinedepartment OR medicinedepartments OR medicinedermatology OR medicinedimed OR medicinedisc OR medicinedivision OR medicinedokkyo OR medicineduke OR medicinedurham OR medicinee OR medicineejacksonville OR medicineemy OR medicineendocrinology OR medicineepidemiology OR medicineerasmus OR medicineeulji OR medicineeuropean OR medicineexch OR medicinefaculty OR medicinefamily OR medicinefariningtonconnecticur OR medicinefederal OR medicinegeneva OR medicinegraduate OR medicinegrand OR medicinehacettepe OR medicinehamad OR medicinehangzhou OR medicinehartford OR medicinehaukeland OR medicinehealth OR medicinehebrew OR medicinehokkaido OR medicinehospital OR medicinehouston OR medicinehubei OR medicinehunan OR medicinehuset OR medicinei OR medicineicams OR medicineilsan OR medicineimperial OR medicinein OR medicineindianapolis OR medicineine OR medicineinfectious OR medicineiningbo OR medicineinneed OR medicineinsight OR medicineinstitut OR medicineinstitute OR medicineinstitutes OR medicineioannina OR medicineisala OR medicineistanbul OR medicinejerusalemisrael OR medicinejiaxing OR medicinejohn OR medicinejohns OR medicinek OR medicinekaohsiung OR medicinekecioren OR medicinekeppel OR medicineking OR medicinekitasato OR medicinekobe OR medicinekocaeli OR medicinekorea OR medicinekyoto OR medicinekyung OR medicinel OR medicinelds OR medicinelerner OR medicinelondon OR medicinelouisiana OR medicineloyola OR medicinem OR medicinema OR medicinemadison OR medicinemakerere OR medicineman OR medicinemas OR medicinemassachusetts OR medicinemayo OR medicinemaywood OR medicinemballage OR medicinemballageoplukning OR medicinemcmaster OR medicinemeaclinical OR medicinemedical OR medicinemedicalsurgical OR medicinememorial OR medicinemen OR medicinement OR medicineminneapolis OR medicinemontplaisir OR medicinemount OR medicinemr OR medicinen OR medicinenagoya OR medicinenanjing OR medicinenational OR medicinenes OR medicinenet OR medicineneuroendocrine OR medicinenew OR medicinenewcastle OR medicinenheden OR medicinenorthwest OR medicinenorthwestern OR medicinenorwegian OR medicinens OR medicinenursing OR medicineo OR medicineodessa OR medicineoffice OR medicineoncologist OR medicineonline OR medicineosaka OR medicineospital OR medicinepaediatric OR medicinepanamericana OR medicinepark OR medicinepediatrics OR medicinepennsylvania OR medicinepet OR medicinepharmacology OR medicinephiladelphia OR medicinephoenix OR medicinephysiology OR medicinepitie OR medicinepittsburgh OR medicinepituitary OR medicineprogram OR medicinepsycho OR medicinepublic OR medicinepubmed OR mediciner OR medicinera OR medicineradboud OR medicineradiologybiomedical OR medicinerar OR medicinere OR medicineresearch OR medicinerhuset OR medicinerigatan OR medicinering OR medicineringen OR medicineringens OR medicineringsfejl OR medicineringsfejlene OR medicineringsproblemer OR medicineringsrutine OR medicinerion OR medicinerlatin OR medicinerna OR medicineroyal OR mediciners OR medicinersang OR medicinersitet OR medicineruddannelsen OR medicinerwth OR medicines OR medicinesahlgrenska OR medicinesaint OR medicinesalt OR medicinesamsung OR medicinesandhealthcareproductsregulatoryagency OR medicinesasang OR medicinesassessment OR medicineschool OR medicinesdevelopment OR medicinesector OR medicinesemmelweis OR medicineseoul OR medicineservice OR medicinesforhumanity OR medicineshanghai OR medicinesincluding OR medicinesir OR medicineslagelse OR medicinesnz OR medicinesodertalje OR medicinespatentpool OR medicinesports OR mediciness OR medicinest OR medicinestanford OR medicinestitute OR medicinestockholm OR medicinesurgical OR medicinesystems OR medicinet OR medicinetai OR medicinetaipei OR medicinetampa OR medicinetask OR medicinetel OR medicinetexas OR medicinethe OR medicinetiken OR medicinetikens OR medicinetisk OR medicinetochigi OR medicinetokyo OR medicinetoo OR medicinetottori OR medicinetri OR medicinetrials OR medicinettuhsctexas OR medicinetufts OR medicinetulane OR medicinetumor OR medicineuc OR medicineumea OR medicineunboxed OR medicineunit OR medicineuniversity OR medicineur OR medicinev OR medicineveterans OR medicinevetsuisse OR medicinevi OR medicinevirginia OR medicinew OR medicinewake OR medicineweill OR medicinewestern OR medicinewestfalische OR medicinewise OR medicinewr OR medicinex OR mediciney OR medicineyale OR medicineyan OR medicineyokohama OR medicinezhejiang OR medicinezhongshan OR (genetic AND therapy) OR genetic OR diabetic OR diabetica OR diabeticae OR diabetical OR diabetically OR diabeticare OR diabeticas OR diabeticcara OR diabeticcomplications OR diabetice OR diabeticek OR diabeticeyes OR diabeticfootcare OR diabeticfoveal OR diabetiche OR diabeticheskaia OR diabeticheski OR diabeticheskie OR diabeticheskii OR diabeticheskikh OR diabeticheskim OR diabeticheskimi OR diabeticheskoe OR diabeticheskogo OR diabeticheskoi OR diabeticheskom OR diabeticheskopi OR diabetichheskoi OR diabetichip OR diabetichiskoi OR diabetichka OR diabetichki OR diabetichnata OR diabetichno OR diabetichskim OR diabetichsskom OR diabetici OR diabeticians OR diabeticiek OR diabeticii OR diabeticilor OR diabeticin OR diabetick OR diabeticka OR diabeticke OR diabetickeho OR diabetickej OR diabetickem OR diabetickom OR diabetickou OR diabeticku OR diabeticky OR diabetickych OR diabetickym OR diabetickymi OR diabeticlike OR diabeticmen OR diabeticne OR diabeticnephropathy OR diabetico OR diabeticorum OR diabeticos OR diabeticpatients OR diabeticpolyneuropathy OR diabeticrats OR diabeticrecipients OR diabeticretinopathy OR diabetics OR diabetics1 OR diabeticsenior OR diabeticsts OR diabetictype OR diabeticul OR diabeticului OR diabeticum OR diabeticus OR diabeticusok OR heart OR (cardio AND renal) AND syndrome OR leopard OR cardiopulmonary OR resuscitation OR blood OR (blood AND pressure) OR blood002 OR blood108 OR blood110 OR bloodagar OR bloodalcohol OR bloodalpha OR bloodaqueous OR bloodartery OR bloodat OR bloodbag OR bloodbank OR bloodbanking OR bloodbankpartners OR bloodbanks OR bloodbanksdelhi OR bloodbased OR bloodbasophil OR bloodbath OR bloodbathed OR bloodbaths OR bloodbeta OR bloodbome OR bloodborn OR bloodborne OR bloodbornes OR bloodbrain OR bloodcapillaries OR bloodcare OR bloodcarrying OR bloodcd4 OR bloodcell OR bloodcells OR bloodcellstorage OR bloodcenter OR bloodcenterofiowa OR bloodcenters OR bloodcentre OR bloodcerebrospinal OR bloodchemical OR bloodchemistry OR bloodchip OR bloodcholesterol OR bloodcirculation OR bloodcleaning OR bloodclearance OR bloodclot OR bloodclots OR bloodclotting OR bloodcoagulation OR bloodcoagulum OR bloodcollections OR bloodcompartment OR bloodcompatibility OR bloodcompatible OR bloodcomponent OR bloodcomponents OR bloodconcentration OR bloodcontaminated OR bloodcount OR bloodcounters OR bloodcounting OR bloodcounts OR bloodculture OR bloodcultures OR bloodcurdling OR bloodderived OR blooddonation OR blooddrawing OR blooded OR bloodedness OR bloodende OR bloodersykdom OR bloodexchange OR bloodexpress OR bloodf OR bloodfed OR bloodfeed OR bloodfeeder OR bloodfeeders OR bloodfeeding OR bloodfeedings OR bloodfeeds OR bloodfilled OR bloodfilling OR bloodfilm OR bloodfilms OR bloodfilter OR bloodfin OR bloodflow OR bloodflower OR bloodflows OR bloodfluke OR bloodflukes OR bloodflux OR bloodform OR bloodforming OR bloodforms OR bloodfree OR bloodfrom OR bloodgas OR bloodgasanalytic OR bloodgases OR bloodgasmonitor OR bloodgasses OR bloodgate OR bloodgen OR bloodglucose OR bloodgood OR bloodgraft OR bloodgroup OR bloodgrouping OR bloodgroupings OR bloodgroups OR bloodh OR bloodhorse OR bloodhound OR bloodhounds OR bloodied OR bloodier OR bloodies OR bloodiest OR bloodin OR bloodiness OR blooding OR bloodinstitute OR bloodisolates OR bloodjournal OR bloodketones OR bloodl OR bloodlactic OR bloodlakes OR bloodlands OR bloodless OR bloodlessly OR bloodlessness OR bloodletings OR bloodletters OR bloodletting OR bloodlevel OR bloodlevels OR bloodlike OR bloodline OR bloodlines OR bloodlink OR bloodlinked OR bloodlipids OR bloodloc OR bloodloss OR bloodlosses OR bloodly OR bloodlymph OR bloodlymphatic OR bloodlymphocytes OR bloodm OR bloodmanagement OR bloodmanager OR bloodmarks OR bloodmax OR bloodmeal OR bloodmeals OR bloodmercury OR bloodmfield OR bloodmobile OR bloodmobiles OR bloodmono OR bloodnerve OR bloodness OR bloodnet OR bloodnfkgar OR bloodning OR bloodninger OR bloodnless OR bloodocular OR bloodof OR bloodomics OR bloodosmotic OR bloodover OR bloodparameters OR bloodpatch OR bloodpatches OR bloodperfused OR bloodperilymph OR bloodpharma OR bloodphobia OR bloodpigments OR bloodplasma OR bloodplatelet OR bloodplatelets OR bloodpmns OR bloodpolymer OR bloodpool OR bloodpooling OR bloodpoolscintigraphy OR bloodpooluntersuchung OR bloodpressure OR bloodpressures OR bloodprints OR bloodproducts OR bloodpump OR bloodpumps OR bloodred OR bloodres OR bloodretinal OR bloodrheological OR bloodrich OR bloodrna OR bloodrood OR sanguinaria OR bloodroot OR bloodrtd OR bloods OR bloodsafety OR bloodsampled OR bloodsamples OR bloodsampling OR bloodsaving OR bloodsaw OR bloodscience OR bloodscreen OR bloodsera OR bloodsermproein OR bloodserum OR bloodservices OR bloodshed OR bloodshedding OR bloodshot OR bloodside OR bloodslide OR bloodslides OR bloodsmear OR bloodsmeared OR bloodsmears OR bloodsource OR bloodspatter OR bloodspilling OR bloodspin OR bloodspinal OR bloodspot OR bloodspots OR bloodspread OR bloodsream OR bloodstage OR bloodstages OR (blood AND stains) OR bloodstain OR bloodstained OR bloodstaines OR bloodstaining OR bloodstasis OR bloodsteam OR bloodstein OR bloodstem OR bloodstemcells OR bloodstill OR bloodstock OR bloodstocks OR bloodstop OR bloodstopper OR bloodstrain OR bloodstrains OR bloodstram OR bloodstrea OR (blood AND circulation) OR bloodstream OR bloodstreamed OR bloodstreamform OR bloodstreaming OR bloodstreamlike OR bloodstreams OR bloodsubstitutes OR bloodsucker OR bloodsuckers OR bloodsucking OR bloodsuckling OR bloodsugar OR bloodsugars OR bloodsupply OR bloodsworth OR bloodsyputum OR bloodsystem OR bloodsystems OR bloodt OR bloodtest OR bloodtesting OR bloodtests OR bloodthirst OR bloodthirstiness OR bloodthirsty OR bloodtight OR bloodtime OR bloodtrack OR bloodtranfusion OR bloodtransfusion OR bloodtransfusions OR bloodtransmitted OR bloodtreatment OR bloodtrykk OR bloodtwig OR bloodtxb OR bloodtype OR bloodtyped OR bloodtyping OR bloodurea OR bloodvalues OR bloodvascular OR bloodvelocity OR bloodvessel OR bloodvessels OR bloodvint OR bloodviscosity OR bloodvolume OR bloodvolumes OR bloodw OR bloodwarmer OR bloodwarmers OR bloodwarming OR bloodwater OR bloodwell OR bloodwings198927 OR bloodwood OR bloodwoods OR bloodwork OR bloodworks OR bloodworm OR bloodworms OR bloodwort OR bloodworth OR bloody OR bloodysputum OR bloodystems OR neoplasms OR hereditary OR ovarian OR cancer OR (early AND detection AND of AND cancer) OR (national AND cancer AND institute) OR (cancer AND vaccines) OR (chemotherapy AND cancer AND regional AND perfusion) OR (cancer AND care AND facilities) OR (american AND cancer AND society) OR (neoplasms AND second AND primary) OR (prostatic AND neoplasms AND castration-resistant) OR urine OR cerebrospinal OR fluid OR (cerebrospinal AND fluid) OR (isolation AND purification) OR psyche OR psychological OR social OR asthma OR cardiorespiratory OR penis OR (lymphatic AND vessels) OR lymphatic OR vessels OR bacteria OR virology OR viral OR microbiology OR fungal OR fungi OR microbial OR toxicity OR (nutritional AND status) OR nutritional OR nutrition OR (nutritional AND sciences) OR food OR nutrients OR (acquired immunodeficiency AND syndrome) OR immunodeficiency OR aids OR hiv OR (heart AND rate) OR heart OR analgesia OR crohn OR diverticulum OR diverticulosis OR methotrexate OR retinaldehyde OR retinal OR retina OR aerosols OR polychlorinated OR biphenyls OR (polychlorinated AND biphenyls) OR pcbs OR mitochondria OR mitochondrial OR animals OR animal OR rats OR rat OR mice OR mouse OR pits OR cattle OR cows OR horses OR horse OR veterinary OR electrocardiography OR electrocardiographic OR ecg OR aorta OR life OR pleura OR pleural OR lung OR lungs OR effusion OR artistic OR sanitary OR hygienic OR metals OR metal OR ocular OR smell OR olfactory OR manipulation OR eating OR eat OR (eating AND disorder) OR (c-reactive AND protein) OR c-reactive OR protein OR angiography OR acupuncture OR (acupuncture AND therapy) OR spleen OR liver OR intestine OR intestines OR bowel OR kidney OR kidneys OR (urinary AND bladder) OR bladder OR (urinary AND tract) OR urinary OR anal OR (administration AND rectal) OR rectal OR (rectal AND administration) OR sexual OR (sexual AND behavior) OR sexually OR treadmill OR foetus OR thyroid OR gland OR thyroid OR trachea OR hearing OR hear OR auscultation OR listening OR (auditory AND perception) OR auditory OR hair OR hairy OR arteries OR artery OR vein OR veins OR vagina OR vaginal OR uterus OR (retroperitoneal AND space) OR retroperitoneal OR embryonal OR embryo OR feces OR fecal OR brains OR epilepsy OR neonatal OR perinatal OR (diabetes AND mellitus) OR diabetes OR mellitus OR (diabetes AND insipidus) OR insipidus OR altitude OR climate OR temperature OR temperatures OR electroencephalography OR eeg OR sleep OR (sleep AND disorder) OR parkinson OR (multiple AND sclerosis) OR sclerosis OR genome OR (genetic AND therapy) OR genetic OR percutaneous OR vaccination OR vaccin OR (influenza AND human) OR influenza OR (human AND influenza) OR laparoscopy OR laparoscopic OR emotions OR emotion OR emotional OR antagonists OR inhibitors OR societies OR society OR opiates OR airway OR breathing OR cerebrum OR cerebral OR brain OR (infant AND newborn) OR newborn OR (newborn AND infant) OR baby OR infant OR endoscopy OR endoscopic OR mental OR mentally OR (substance-related AND disorders) OR substance-related OR abuse OR fibroblasts OR fibroblast OR (radioisotope AND renography) OR radioisotope OR renography OR pharmacokinetic OR molecular

OR laparotomy OR prophylactic OR sarcopenia OR sarcopenie OR insurance OR mortality OR dermatologic OR baroreflex OR baroreflexes OR schizophrenia OR (neuro AND cognitive) OR cognitive OR burns OR burn OR groin OR inguinal OR mortal OR tongue OR nose OR lip OR lips OR cheek OR cheeks OR radiation OR (electromagnetic radiation) OR depressive OR (depressive AND disorder) OR (anxiety AND disorder) OR ulcera OR psoriasis OR psoriatic OR puva OR glucosamine OR insulin OR myope OR myopy OR presbyopes OR bicycle OR nursing OR (breast AND feeding) OR feeding OR infection OR (electrical AND stimulation) OR fes OR behaviour OR behavior OR (down AND syndrome) OR (job AND satisfaction) OR (osteo AND densitometry OR memory OR diet OR diets OR respiratory OR respiration OR (cell AND respiration) OR aspiration OR inhalation OR inspiration OR exhalation OR expiration OR saturation OR embalming OR (socio AND economic) OR fall OR calls OR esthetics OR visceral OR spermatozoa OR sperm OR myocardium OR myocardial OR infarction OR infarct OR (mutagenesis AND insertional) OR mutagenesis OR insertional OR (insertional AND mutagenesis) OR insertion OR (6-min AND walk AND test) OR fetus OR fetal OR (heart AND failure) OR dentin OR dental OR tonsillectomy OR caregiving OR ectopic OR dogs OR dog OR canines OR intragastric OR stomach OR gastric OR cryotherapy OR (pressure AND ulcer) OR ulcer OR decubitus OR mammography OR hasabstract OR cardior* OR cardiov* OR cardiop* OR cardioa* OR cardioc* OR cardiod* OR cardioe* OR cardiof* OR cardiom* OR cardios* OR anxiety OR (bone AND mineral AND density) OR malocclusion OR endocrinology OR airbag OR motorcycle OR language OR (laser AND therapy) OR shops OR antibiotic OR micronized OR (vestibular AND evoked) OR (vestibule AND spinal) OR (cosmetic AND surgery) OR hirsutism OR copd OR cervicovaginal OR mucus OR clubhead OR sociocultural OR (growth AND hormone AND treatment) OR (attention AND deficit AND disorder) OR horman OR hormonal OR (hyperactivity AND disorder) OR incision OR (vibration AND training) OR (berg AND balance AND scale) OR (wound AND therapy) OR wound OR prescriptions OR herbal OR (traditional AND medicine) OR (Chinese AND medicine) OR bio-effects)

Similar search strings were used for the digital data bases of Cochrane (1940-2017), Medline (PubMed) (1950-2017), Embase (1974-2017), CISDOC (1901-2017), ScienceDirect (1997-2017) and CINAHL (1977-2017) and available at the corresponding author.

Supplement table 2:

Checklist for the assessment of clinical usability of posture assessment instruments.

1. Readability/comprehensibility:
   a. Good: **2 points**
   b. Reasonable / moderate: **1 point**
   c. Bad: **-1 point**
   d. Unclear: **0 points**
2. Execution time test:
   a. < 2 min.: **2 points**
   b. 2-5 min.: **1 point**
   c. > 5 min.: **-1 point**
   d. Unclear: **0 points**
3. Costs equipment:
   a. < €250: **2 points**
   b. € 250-500: **1 point**
   c. € 500-2000: **-1 point**
   d. > € 2000: **-2 points**
   e. Unclear: **0 points**

4. Physical supplies:
   a. Supplies are standard, easy to apply and is not annoying space in standard research area: **2 points**
   b. Supplies are non-standard, not easy to apply and / or requires more space (but applicable) in standard research area: **1 point**
   c. Supplies are hard to make, hard to get, requires much space or annoying to be compiled in separate research area: **-1 point**
   d. Unclear or inadequate description regarding supplies: **0 points**

5. Interpretation load:
   a. Easy to interpret, unambiguous score or one answer, execution time < 2 min: **2 points**
   b. Fairly easy to interpret, composite score in subdomains or brief description or tag, execution time < 5 min: **1 point**
   c. Interpretation difficult or time consuming, execution time > 5 min: **-1 point**
   d. Unclear or inadequate description regarding interpretation load: **0 points**
6. Patient friendliness:
   a. Patient friendly: **1 point**
   b. Patient unfriendly (ie. completely undress): **-1 point**
   c. Unclear: **0 points**
7. Level of education evaluator and test interpretation:
   a. Short, easy and unambiguous, < 10 min. training: **2 points**
   b. Longer or interpretation requires some specific knowledge, < 1 hour training: **1 points**
   c. Difficult or for interpretation is a long learning process required, > 1 hour: **-1 point**
   d. Unclear: **0 points**

Supplement table 3: Table with the characteristics of the studies.

Author (ref.no.)	Study properties	Instrument name	Country	Study setting	Target population	Study population (n)	Mean age/ SD/range (years)	Gender (M / F)
Ackermann (21)	Inter-rater reliability	-	Australia	Laboratory	Musicians	15 violinists, 15 violinists with pain	Unclear	Unclear
Bao (52)	Criterion validity	Rapid Upper Limb Assessment (RULA) (Event Based Method versus Time Based Method)	USA	Work environment	Working population	733	Male 36.6/-/- Female 42/-/-	383/350
Bao (52)	Inter-rater reliability	Rapid Upper Limb Assessment (RULA)	USA	Work environment	Working population	4	Unclear	Unclear
Barker (54)	Intra-rater reliability Inter-rater reliability	Videopoint	USA	Laboratory?	General population	10 students/ 3 raters	Unclear	Unclear
Barks (55)	Content-validity Construct-validity Test-retest	Seated Posture Scale	USA	Veterians home/laboratory	Older wheelchair users	6 experts/ 49 veterians (>63 yrs)/ 3 dummies	-/-/60-99	48/1
Blanken (7)	Intra-rater reliability Inter-rater reliability	-	The Netherlands	Unclear	Musicians	24	20/-/17-29	11/13
Brink (56)	Test-retest reliability Measurement error	3D-Posture Analysis Instrument (3D-PAT)	South-Africa	Laboratory	General population	1 mannequin 24 persons	16,1/-/-	10/14
de Bruijn (57)	Intra-rater reliability Inter-rater reliability	Ovako Working Posture Analyzing System (OWAS)	The Netherlands	Work environment	Nurses	Intra-rater reliability (≤24hrs):32 Intra-rater reliability (4 weeks): 31 Inter-rater reliability: 45	Unclear	Unclear
Bunkan (58)	Inter-rater reliability	Comprehensive Body Examination	Norway	Hospital & rehabilitation setting	General population	25	34/-/20-57	6/19

Author (ref.no.)	Study properties	Instrument name	Country	Study setting	Target population	Study population (n)	Mean age/ SD/range (years)	Gender (M / F)
Carr (59)	Inter-rater reliability	Initial Pictorial Instrument Modified Pictorial Instrument	United Kingdom	Hospital setting	Stroke patients	57	63.8/-/33-86	35/22
David (60)	Intra-rater reliability Inter-rater reliability Criterion validity Usability	Quick exposure check (QEC)	United Kingdom	Work environment	Working population	Unclear	Unclear	Unclear
Dockrell (61)	Intra-rater reliability Inter-rater reliability	Rapid Upper Limb Assessment (RULA)	Ireland	Primary school	Children	24	-/-/4-12	12/12
Dunk (62)	Intra-rater reliability Inter-rater reliability	-	Canada	Laboratory	General population	14	21.8/-/-	7/7
Dunk (63)	Test-retest reliability Measurement error	-	Canada	Laboratory	General population	20	22.2/-/-	10/10
Eriksson (64)	Inter-rater reliability	Resource-oriented Physiotherapeutic Examination	Sweden	Hospital	General population	19	46/7/-	3/16
Fedorak (29)	Intra-rater reliability Inter-rater reliability	-	Canada	unclear	General population	36	41/-/18-73	18/18
Ferreira (8)	Intra-rater reliability Inter-rater reliability Measurement error (Criterion val.)	Postural Assessment Software (PAS) / SAPO	Brazil	Laboratory	General population	22	Unclear	Unclear
Fortin (65)	Intra-rater reliability Inter-rater reliability Measurement error	-	Canada	Laboratory	People with idiopathic scoliosis	70	15.7/-/10-20	10/60

Supplement table 3 continued

Author (ref.no.)	Study properties	Instrument name	Country	Study setting	Target population	Study population (n)	Mean age/ SD/range (years)	Gender (M / F)
Furlanetto (66)	Intra-rater reliability Inter-rater reliability Test-retest Measurement error Concurrent-Validity	-	UK	Laboratory	General population	16 54	23.7/3.6/- 45.4/18.1	8/8 25/29
Harrison (67)	Inter-rater reliability	Unclear	USA	Laboratory	General population	41	-/-/20-45	11/30
Haugstad (68)	Inter-rater reliability (Discriminative validity)	Standardized Mensendieck physiotherapy Test (SMT)	Norway	Laboratory	General population	30 (Healthy: 15; Chronic pelvic pain subject:15)	-/-/20-50	0/30
Hignett (69)	Inter-rater reliability	Rapid Entire Body Assessment (REBA)	United Kingdom	Healthcare-, manufacturing- and electricity-industries.	Working population	600	Unclear	Unclear
Karim (70)	Inter-rater reliability	Unclear	USA	Laboratory	Dancers	30	24/-/18-32	0/30
Kvale (71)	Discriminative validity Internal consistency	Global Body Examination – New Posture Domain	Norway	Laboratory	General population	130 (Healthy: 32, local pain: 32, widespread pain: 32, psychosis: 34)	38.7/12.5/-	43/89
Liebregts (72)	Inter-rater reliability Concurrent-validity	Rapid Office Strain Assessment (ROSA)	Canada	Computer work stations	Working population	23	Unclear	11/12
McAlpine (73)	Intra-rater reliability Inter-rater reliability	Middleborough Integrated Digital Assessment System (MIDAS)	United Kingdom	Laboratory	Healthy young people	25	24.7/-/ 23-28	5/20
McEvoy (74)	Test-retest reliability	Unclear	Australia	Primary school	Primary school students	38	-/-/5-12	23/15

Supplement table 3 continued

Author (ref.no.)	Study properties	Instrument name	Country	Study setting	Target population	Study population (n)	Mean age/ SD/range (years)	Gender (M / F)
Menendez (75)	Concurrent-validity	Modified rapid upper limb assessment (mRULA) & University of California Computer Use Checklist (UCCUC)	USA	University	Students	30	-/-/≥18	15/15
Motamedzade (75)	Concurrent validity	Rapid Upper Limb Assessment (RULA) & Quick Exposure Check (QEC)	Iran	Engine oil company	Working population	40	Unclear	Unclear
van Niekerk (77)	Test-retest reliability Measurement error Concurrent-validity	Photographic Posture Analysis Method (PPAM)	South-Africa	Laboratory	Adolescent	39	-/-/15-16	19/20
Normand (77)	Intra-rater reliability Inter-rater reliability Measurement error	PosturePrint	Canada	Laboratory	General population	40	24.4/1.9)/-	10/30
Paquet (79)	Concurrent validity	Posture, activities, instruments and handling(PATH; modified from OWAS) & Simplified PATH	USA	Laboratory	Working population	5	Unclear	5/0
Paul (80)	Criterion validity	2D photography	The Netherlands	Laboratory	General population	6	26.5/2.8/-	0/6
Pausic (81)	Intra-rater reliability Measurement error Concurrent validity	Posture-image Analyzer Software & University of Texas Health Science Centre San Antonio Image Instrument (UTHSCSAIT)	Croatia	Elementary school	Elementary school Students	273	-/-/10-13	Unclear

Supplement table 3 continued

Author (ref.no.)	Study properties	Instrument name	Country	Study setting	Target population	Study population (n)	Mean age/ SD/range (years)	Gender (M / F)
Perry (82)	Intra-rater reliability (digitization) Inter-rater reliability (examiners) Measurement error	Unclear	Australia	Laboratory	Adolescents	22	-/-/13-17	11/11
Pownall (83)	Test-retest reliability Measurement error	Unclear	New Zealand	Laboratory	General population	11	29.6 /10.4/-	11/0
Rodby-Bosquet (84)	Inter-rater reliability Construct-validity	Posture and Postural Ability Scale (PPAS)	Sweden & Iceland	Rehabilitation centre	Adults with cerebral palsy	30	-/19-22	15/15
Ruivo (85)	Intra-rater reliability Inter-rater reliability	Postural Assessment Software (PAS)	Portugal	Secondary School	Adolescents	17	15.8/15-17	3/14
Sanchez (86)	Inter-rater reliability Concurrent-validity	DARTFISH	UK	Unclear/ laboratory	General population	12	27.9/3.5/-	6/6
Wilbanks (87)	Intra-rater reliability Inter-rater reliability Test-retest reliability	-	USA	Laboratory	General population	10	-/-/19-60	Unclear
Zonnenberg (88)	Intra-rater reliability Inter-rater reliability	-	The Netherlands	Laboratory	General population	18	Unclear	Unclear

Supplement table 4: QUADAS-II: Methodological quality of the included studies.

Study (ref.)	RISK OF BIAS				APPLICABILITY CONCERNS		
	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD	FLOW AND TIMING	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD
Ackermann (21)	?	?	--	?	☺	☺	--
Bao (52)	?	?	--	--	?	--	--
Bao (53)	?	☹	--	--	☹	☹	--
Barker (54)	?	?	--	?	☺	☺	--
Barks (55)	☺	?	--	☺	☺	?	--
Blanken (7)	?	?	--	☺	☺	☺	--
Brink (56)	--	☹	--	--	☹	☹	--
de Bruijn (57)	?	☺	--	?	?	☺	--
Bunkan (58)	?	☹	--	--	☹	☺	--
Carr (59)	?	☺	--	☺	☺	☺	--
David (60)	?	?	--	?	?	☺	--
Dockrell (61)	?	?	--	?	☺	☺	--
Dunk (62)	☹	☺	--	☺	☹	☹	--
Dunk (63)	?	?	--	☺	☺	☺	--
Eriksson (64)	☹	?	--	☺	?	☺	--
Fedorak (29)	☹	☹	--	☺	☺	☺	--
Ferreira (8)	?	?	--	☺	?	?	--
Fortin (65)	☺	☺	--	☺	☺	☺	--
Furlanetto (66)	☺	☺	?	☺	☺	☺	☺
Harrison (67)	?	☺	--	?	☺	☺	--
Haugstad (68)	?	☹	--	?	☹	☺	--
Hignett (69)	?	?	--	?	?	☺	--
Karim (70)	☺	☺	--	☺	☺	☺	--
Kvale (71)	?	☺	☺	☺	☺	☺	☺
Liebregts (72)	☺	☺	?	☺	?	☺	?
McAlpine (73)	☺	☹	--	☺	☺	☹	--
McEvoy (74)	?	☺	--	☺	☺	☺	--
Menendez (75)	☺	?	?	?	☺	☺	☺
Motamedzade (76)	?	?	?	?	☺	☺	☺
van Niekerk (77)	☺	?	?	☺	☺	☺	☺
Normand (78)	☺	☺	--	☺	☺	☺	--
Paquet (79)	☹	?	?	☺	☹	?	?
Paul (80)	☺	☺	☺	☺	☺	☺	☺
Pausic (81)	☹	☹	?	☺	☺	☺	☺
Perry (82)	☺	☺	☺	☺	☺	?	--
Pownall (83)	☺	☹	--	☺	☹	☹	--
Rodby-Bousquet (84)	☺	☹	--	☺	☺	☺	--
Ruivo (85)	☺	☺	--	☺	☺	☺	☺
Sanchez (86)	?	☺	☺	☺	☺	☺	☺
Wilbanks (87)	?	?	--	☺	☺	☺	--
Zonnenberg (88)	?	?	--	☺	?	☺	--

☹ = High risk of bias or high concerns about applicability; ☺ = Low risk of bias or low concerns about applicability; ? = Unclear risk of bias or unclear concerns about applicability; - = not applicable

Supplement table 5

Supplement table 5: Table with the characteristics of the assessment instruments.

	Instrument or observation method	Author (ref.)	Name of instrument	Type of instrument	Observation goal #	Body domains (number)	Type of scale	Number of items[1]	Range of scores
Body measurement: Direct method	Measurements with digital optic sensors of predefined anatomical landmarks • MIDAS	McAlpine (73)	The Middleborough Integrated Digital Assessment System (MIDAS)	Observation with Optic sensors	Standing and sitting posture	Head & trunk (2)	Continuous	3/12	Angles (degrees) & distances (mm)
	Manual analysis with ruler, goniometer & reference line	Harrison (67)	-	Ruler, goniometer & reference line	Standing posture	Head & upper extremity (2)	Ratio	2/3	Angles (degrees) & distances (mm)
Body measurement: Indirect method	Manual analysis of digital photographs • PAS • UTHSCSAIT	Pausic (81)	Posture-Image Analyzer Software (PAS) University of Texas Health Science Centre, San Antonio Image Instrument (UTHSCSAIT)	Photo camera, anatomical reflective markers & reference line	Standing posture	Head, lower extremities & trunk (3)	Continuous	1/2	Angles (degrees) & distances (mm)
	Analog analysis of analog photographs	Zonnenberg (88)	-	Photo camera & gridline reference	Standing posture	Head & trunk (2)	Continuous	2/8	Distances (mm)
Visual assessment: Direct method	Instantaneous visual analysis with pre-defined posture category system	Blanken (7)	-	Visual observation	Standing and sitting music making posture	Head, upper & lower extremities and trunk (4)	Nominal	1/7	3-12
	Instantaneous visual analysis with pre-defined posture category system • CBE	Bunkan (58)	Comprehensive Body Examination (CBE)	Visual observation	Upright, supine and bent forward sitting and standing posture	Head & trunk (2)	Interval & dichotomous	12	1-7 / 1-2
	Instantaneous visual analysis with pre-defined posture category system • IPI • MPI	Carr (59)	Initial Pictorial Instrument (IPI) Modified Pictorial Instrument (MPI)	Visual observation	Sitting, supine and lateral lying	Head, upper & lower extremities and trunk (4)	Ordinal	2/8	4-6
	Instantaneous visual analysis with pre-defined posture category system • RPE	Eriksson (64)	Resource –oriented Physiotherapeutic Examination (RPE)	Visual observation with gridline reference	Standing and supine posture	Head, upper & lower extremities and trunk (3/4)	Nominal & ordinal	Upright 1/4 Supine 1/2	Upright 2/7 supine 3/5
	Instantaneous visual analysis with pre-defined posture category system • SMT	Haugstad (68)	Standardized Mensendieck physiotherapy Test (SMT)	Visual observation	Standing posture	Head, upper & lower extremities, trunk and centre of gravity (5)	Interval	1/3	4

Supplement table 5 continued

	Instrument or observation method	Author (ref.)	Name of instrument	Type of instrument	Observation goal #	Body domains (number)	Type of scale	Number of items[1]	Range of scores
Visual assessment: Direct method	Instantaneous visual analysis with pre-defined posture category system • MS	Karim (70)	Musculoskeletal Screen (MS)	Visual observation	Standing posture	Lower extremities & trunk (2)	Nominal & dichotomous	4/5	2
	Instantaneous visual analysis with pre-defined posture category system • GBE	Kvale (71)	Global body Examination (GBE)	Visual observation	Standing and supine posture	Head, upper & lower extremities, trunk and centre of gravity (5)	Ordinal & dichotomous	1/4	2-3
	Instantaneous visual analysis with pre-defined posture category system • REBA • QEC	Motamedzade (76)	Rapid Entire Body Assessment (REBA) - Quick Exposure Check (QEC)	Visual observation	Standing and sitting working posture	Head, upper & lower extremities and trunk (4)	Ordinal	1/4 / 3	2-7 / 2-3
	Instantaneous visual analysis with pre-defined posture category system • PATH • sPATH	Paquet (79)	Posture, Activities, Instruments and Handling (PATH) Simplified PATH (sPATH)	Visual observation	Standing working posture	Upper & lower extremities and trunk (3)	Ordinal	3	3-7
	Instantaneous visual analysis with pre-defined posture category system • mRULA • UCCUC	Menendez (75)	Modified- Rapid Upper Limb Assessment (mRULA) & University California Computer Use Checklist (UCCUC)	Visual observation	Standing and sitting posture	Head, upper & lower extremities and trunk (4)	Ordinal	1/3 / 1/6	4-8 / 3-7
Visual assessment: Indirect method - photograph	Visual analysis of photographs with pre-defined posture category system • SPS	Barks (55)	Seated Posture Scale (SPS)	Visual observation	Sitting posture	Head, upper & lower extremities and trunk (4)	Dichotomous	23	0-1
	Visual analysis of photographs with pre-defined posture category system • OWAS	de Bruijn (57)	Ovako Working Posture Analyzing System (OWAS)	Photo camera	Standing and sitting working posture	Head, upper & lower extremities and trunk (4)	Nominal	1	3/7
	Visual analysis of photographs with pre-defined posture category system	Fedorak (29)	-	Photo camera	Standing posture	Head & trunk (2)	Interval	1	3

Supplement table 5 continued

	Instrument or observation method	Author (ref.)	Name of instrument	Type of instrument	Observation goal #	Body domains (number)	Type of scale	Number of items[1]	Range of scores
Visual assessment: Indirect method - photograph	Visual analysis of photographs with pre-defined posture category system • REBA	Hignett (69)	Rapid Entire Body Assessment (REBA)	Photo camera	Standing and sitting posture	Head, upper & lower extremities and trunk (4)	Ordinal	1/3	2-7
	Visual analysis of photographs with pre-defined posture category system • ROSA	Liebregts (72)	Rapid Office Strain Assessment (ROSA)	Photo camera	Sitting posture	Head, upper & lower extremities and trunk (4)	Ordinal	8	1-6
	Visual analysis of photographs with pre-defined posture category system • PPAS	Rodby-Bousquet (84)	Posture and Postural Ability Scale (PPAS)	2 Photo cameras	Standing, sitting, prone or supine posture	Head, upper & lower extremities, trunk and centre of gravity (5)	Ordinal (composed of dichotomous scales)	1-4	0-7
	Visual analysis of photographs with pre-defined posture category system system	Wilbanks (87)	-	Photo camera	Standing posture	Head and trunk (2)	Continuous	4	Angles (degrees) & distances (mm)?
Visual assessment: Indirect - video	Visual analysis of video with pre-defined posture category system • mRULA EBM • mRULA TBM	Bao (52)	Modified RULA: - mRULA EBM - mRULA TBM	Two digital video cameras	EBM: most common and worst standing and sitting posture TBM: calculation both postures at random in time	Head, upper extremity & trunk (3)	Ordinal	3/14	3-5
	Visual analysis of video with pre-defined posture category system	Bao (53)	-	Two digital video cameras	Standing and sitting posture	Head, upper extremity & trunk (3)	Interval	3	1-5
	Visual analysis of video with pre-defined posture category system • QEC • $QEC_{modified}$	David (60)	Quick Exposure Check (QEC)	Video camera	Standing posture	Head, upper extremities & trunk (3)	Ordinal	1/2	3-5
	Visual analysis of video without categorical system • Visual Analog Scale (VAS)	Ackermann (21)	-	Video camera	Standing posture	Head, trunk, centre of gravity (3)	Continuous	1-2	Distances (mm)
	Visual analysis of video with pre-defined posture category system • RULA	Dockrell (61)	Rapid Upper Limb Assessment (RULA)	Video camera	Standing and/or sitting computing posture	Head, upper & lower extremities and trunk (4)	Ordinal	1/4	2/7

Supplement table 5 continued

	Instrument or observation method	Author (ref.)	Name of instrument	Type of instrument	Observation goal #	Body domains (number)	Type of scale	Number of items[1]	Range of scores
Digital assessment	Digital analysis of digital photographs without categorical system • Videopoint	Barker (54)	Videopoint	Digital photo camera	Sagittal plane standing or sitting posture	Head, upper extremity & trunk (3)	Continuous	1/2	Angles (degrees)
	Digital analysis of digital photographs without categorical system • 3D-PAT	Brink (56)	3D-Posture Analysis Instrument (3D-PAT)	Five digital photo cameras	Standing and/or sitting computing posture	Head & trunk (2)	Continuous	2/7	Angles (degrees)
	Digital analysis of digital photographs without categorical system	Dunk (62)	-	Photo camera, anatomical landmarks & vertical ref line	Standing posture	Head & trunk (2)	Continuous	1/2	Angles (degrees)
	Digital analysis of digital photographs without categorical system	Dunk (63)	-	Photo camera, anatomical landmarks and vertical ref line	Standing posture	Head & trunk (2)	Continuous	2/4	Angles (degrees)
	Digital analysis of digital photographs without categorical system • PAS	Ferreira (8)	Postural Assessment Software (PAS)	Two photo cameras & reference line	Standing posture	Head, upper & lower extremities and trunk (4)	Continuous	2/16	Angles (degrees) & distances (mm)
	Digital analysis of digital photographs without categorical system	Fortin (65)	-	Two photo camera & anatomical landmarks	Standing posture	Head, upper & lower extremities and trunk (4)	Continuous	4/11	Angles (degrees) & distances (mm)
	Digital analysis of digital photographs without categorical system • DIPA	Furlanetto (66)	Digital Image-Based Posture Assessment (DIPA)	photo camera, anatomical landmarks	Standing posture	lower extremities and trunk (2)	Continuous	1/5	Angles (degrees) & distances
	Digital analysis of analog photographs (negatives) without categorical system	McEvoy (74)	-	Photo camera & reference line	Standing posture	Head, lower extremities & trunk (3)	Continuous	1/3	Angles (degrees)
	Digital analysis of digital photographs without categorical system • PPAM	van Niekerk (77)	Photographic Posture Analysis Method (PPAM)	Photo camera	Sitting working posture	Head, upper extremities & trunk (3)	Continuous	1/2	Angles (degrees) & distances (mm)
	Digital analysis of digital photographs without categorical system • PosturePrint	Normand (78)	PosturePrint	Three photo cameras & gridline reference lines	Standing posture	Head, upper & lower extremities and trunk (4)	Continuous	5/11	Angles (degrees) & distances (mm)
	Digital analysis of analog photographs without categorical system	Paul (80)		2 photo cameras, anatomical landmarks & reference line	Standing posture	Head, upper & lower extremities and trunk (4)	Continuous	1/3	Angles (degrees)

Supplement table 5 continued

	Instrument or observation method	Author (ref.)	Name of instrument	Type of instrument	Observation goal #	Body domains (number)	Type of scale	Number of items[1]	Range of scores
Digital assessment	Digital analysis of digital photographs without categorical system • PIAS • UTHSCSAIT	Pausic (81)	Posture-Image Analyzer Software (PIAS) University of Texas Health Science Centre, San Antonio Image Instrument (UTHSCSAIT)	Photo camera, anatomical reflective markers & reference line	Standing posture	Head, lower extremities & trunk (3)	Continuous	1/2	Angles (degrees) & distances (mm)
	Digital analysis of digital photographs without categorical system	Perry (82)	-	Photo camera, anatomical reflective markers & reference line	Standing, sitting or slumped posture	Head, trunk & centre of gravity (3)	Continuous	1/6	Angles (degrees) & distances (mm)
	Digital analysis of digital video stills without categorical system	Pownall (83)		2 Video cameras, anatomical landmarks & reference line	Standing or sitting posture	Head, upper & lower extremities and trunk (4)	Continuous	3/4	Angles (degrees) & distances (mm
	Digital analysis of digital photographs without categorical system • PAS/SAPO	Ruivo (85)	Posture Assessment System (PAS/SAPO)	Photo camera, anatomical reflective markers	Standing posture	Head, upper extremities (2)	Continuous	1/2	Angles (degrees)
	Digital analysis of digital photographs without categorical system	Sanchez (86)	-	Video system	Sitting posture	Head, upper extremities & trunk (3)	Continuous	1/3	Angles (degrees)

[1] Because the assessment instruments consists ≥ 2 domains; the minimal and maximal number of items per domain-scale.

Observation goal: Standing and sitting posture= data presentation together; Standing or sitting posture= data presentation separated per posture; Standing and/or sitting posture= Unclear if standing posture is included in computing working posture.

Supplement table 6

Supplement table 6: Table of the clinimetric characteristics of the assessment instruments.

	Instrument or observation method	Author (ref.)	Intra-rater reliability	Inter-rater reliability	SEM	Test-retest	Construct validity	Criterion/ Concurrent validity	Content validity
Body measurement: Direct method	Middleborough Integrated Digital Assessment System (MIDAS)	McAlpine (73)	ICC_{mean}: 0.967 ICC: 0.92 – 0.99	ICC_{mean}: 0.97 ICC: 0.90-0.98	-	-	-	-	-
	Manual analysis with ruler, goniometer & reference line	Harrison (67)	-	ICC: 0.34-0.91	-	-	-	-	-
Body measurement: Indirect method	Manual analysis of digital photographs	Pausic (81)	UTHSCSAIT software: ICC:0.80-0.91	-	UTHSCSAIT software: SEM: 0.59-0.91	-	Inter method agreement (PIAS – UTHSCSAIT) ICC: 0.992-0.998	-	-
	Analog analysis of analog photographs	Zonnenberg (88)	ICC: 0.7324-1.000 Pearson product-moment correlation 0.7523-1.000	ICC: 0.66-1.00 Pearson product-moment correlation: 0.73-1.00	-	-	-	-	-
Visual assessment: Direct method	Instantaneous visual analysis with pre-defined posture category system	Blanken (7)	ICC: -0.20-1.00 K_{mean}: 0.51-0.59	ICC : -0.10-0.80 K_{mean}: 0.25-0.48	-	-	-	-	-
	Instantaneous visual analysis with pre-defined posture category system • CBE	Bunkan (58)	-	ICC: 0.79-0.99 ICC_{mean}: 0.94	-	-	Internal consistency Cronbachs α: 0.61-0.96 Cronbachs α_{Median}:0.83	-	-
	Instantaneous visual analysis with pre-defined posture category system • IPT • MPT	Carr (59)	-	Posture obs initial K >0.6: 98.5% PoA >70%: 100% Posture obs modified K >0.6: 15.8% PoA >70%: 52.6%	-	-	-	-	-
	Instantaneous visual analysis with pre-defined posture category system • RPE	Eriksson (64)	-	K_{mean}: 0.55 K: 0.23-0.94 (95% CI) PoA: 75%	-	-	-	-	-
	Instantaneous visual analysis with pre-defined posture category system • SMT	Haugstad (68)	-	Standing ICC_{mean}: 0.89 ICC: 0.83-0.91 Sitting ICC_{mean}: 0.93 ICC: 0.89-0.97	-	-	-	-	-
	Instantaneous visual analysis with pre-defined posture category system • MS	Karim (70)	-	PoA: 43.3%-80%	-	-	-	-	-

Supplement table 6 continued

	Instrument or observation method	Author (ref.)	Intra-rater reliability	Inter-rater reliability	SEM	Test-retest	Construct validity	Criterion/ Concurrent validity	Content validity
	Instantaneous visual analysis with pre-defined posture category system • GBE	Kvale (71)	-	-	-	-	Subscale peripheral flexion: Inter item correlation 0.30-0.44 Cronbachs α: 0.68 Subcale pelvic tilt: Inter item correlation: 0.48-0.70 Cronbachs α: 0.78 Subscale body axis: Inter item correlation: 0.27-0.59 Cronbachs α: 0.63 Inter correlation between subscales 0.08-0.31 Inter item correlation Cronbachs α: 0.68 AUC 0.82-0.96	CBE AUC: 0.86-0.97 GPE52 AUC: 0.69-084 NPD AUC: 0.82-0.96	-
Visual assessment: Direct method	Instantaneous visual analysis with pre-defined posture category system • REBA • QEC	Motamedzade (76)	-	-	-	-	-	P value of five working departments 0.157-0.800 QEC – REBA correlation action levels: 0.89 Correlation final score: R: 0,73	-
	Instantaneous visual analysis with pre-defined posture category system • PATH • sPATH	Paquet (79)	-	-	-	-	-	Shoulder: p=0.54; Knee p=0.02; Trunk p=0.01. Agreement shoulders PoA: 93%-95% K: 0.74-0.80 Knee PoA: 81%-90% K: 0.75-0.87 Trunk PoA: 68%-75% K: 0.51-0.60	-

Supplement table 6 continued

	Instrument or observation method	Author (ref.)	Intra-rater reliability	Inter-rater reliability	SEM	Test-retest	Construct validity	Criterion/ Concurrent validity	Content validity
Visual assessment: Direct method	Instantaneous visual analysis with pre-defined posture category system • mRULA • UCCUC	Menendez (75)	-	-	-	-	-	mRULA: Right side: P=0.93 Pearson's R^2: 0.09 Left side: P=0.87 Pearson's R^2: 0.018 mRULA OR posture assessment: 0.75 (0.5-1.1) UCCUC OR posture assessment: 1.3 (1.0-1.6)	-
Visual assessment: Indirect method - photograph	Visual analysis of photographs with pre-defined posture category system	Barks (55)	-	-	-	ICC= 0.89-0.99	r with MAS= - 0.44 r with BI= 0.46/ p<0.05 r with VDS= 0.22/ p= 0.12	-	0.74 (0.17-1.00) 10/24 items perfect score
	Visual analysis of photographs with pre-defined posture category system • OWAS	de Bruijn (57)	Short: K_{mean}: 0.91 PoA: 92% Long: K_{mean}: 0.98 Poa: 90%	K_{mean}: 0.87 PoA: 89%	-	-	-	-	-
	Visual analysis of photographs with pre-defined posture category system	Fedorak (29)	K_{mean}: 0.50 K: 0.02-0.98 (95% CI)	K_{mean}:0.16 K: 0.00-0.48 (95% CI)	-	-	-	-	-
	Visual analysis of photographs with pre-defined posture category system • REBA	Hignett (69)	-	PoA: 62%-84%	-	-	-	-	-
	Visual analysis of photographs with pre-defined posture category system • ROSA	Liebregts (72)	-	0.67-0.86	-	-	3.1-47	-	-

Supplement table 6 continued

	Instrument or observation method	Author (ref.)	Intra-rater reliability	Inter-rater reliability	SEM	Test-retest	Construct validity	Criterion/ Concurrent validity	Content validity
Visual assessment: Indirect method - photograph	Visual analysis of photographs with pre-defined posture category system ROSA	Rodby-Bousquet (84)	$K_{weighted}$: Sitting: 0.87-0.91 Standing: 0.95		-	-	Internal consistency (Cronbach α): Sitting: 0.96-0.97 Standing: 0.96 Item-total correlation: Sitting: 0.60-0.84 Standing:0.81-0.82 GMFCS (P-value): Sitting: <0.001-0.019 Standing: <0.001	-	-
	Visual analysis of photographs with pre-defined posture category system	Wilbanks (87)	0.87-0.98 / p <0.05		?	0.62-0.80 / p < 0.05	-	P=0.07; $SD_{digital\ photo}$ < $SD_{manually}$ (no data)	-
Visual assessment: Indirect method - video	Visual analysis of video with pre-defined posture category system • mRULA EBM • mRULA TBM	Bao (52)	-	-	-	-	-	K most common posture: 0.00-0.24 K worst posture: -0.01-0.27	-
	Visual analysis of video with pre-defined posture category system	Bao (53)	-	ICC_{mean}: 0.14-0.77 PoA: 45%-95%	3.8^0-27.8^0	-	-	-	-
	Visual analysis of video with pre-defined posture category system • QEC • $QEC_{modified}$	David (60)	QEC: K: 0.45-0.53 Spearman: 0.45-0.69 PoA: 66.7%-76.7%	QEC: K: 0.00-0.47 PoA: 64.7%-80.2% $QEC_{modified}$: K:0.6 Spearman 0.79-0.98 PoA: 83%-100%	-	-	-	SIMI versus practicioners PoA: 72.%-87.5%	-
	Visual analysis of video without categorical system • VAS	Ackermann (21)	ICC: 0.2-0.65	ICC: -0.15-0.60	-	-	-	-	-
	Visual analysis of video with pre-defined posture category system	Dockrell (61)	Grand scores: ICC: -0.23-0.93 Action level: ICC -0.69-0.94	Grand scores: ICC 0.01-0.86 Action levels: ICC 0.09-0.85	-	-	-	-	-

Supplement table 6 continued

	Instrument or observation method	Author (ref.)	Intra-rater reliability	Inter-rater reliability	SEM	Test-retest	Construct validity	Criterion/ Concurrent validity	Content validity
Digital assessment	Digital analysis of digital photographs without categorical system • Videopoint	Barker (54)	ICC: 0.35-0.95	ICC: 0.32-0.86	-	-	-	-	-
	Digital analysis of digital photographs without categorical system • 3D-PAT	Brink (56)	-	-	-	Mannequin CCC: 0.99 Students ICC: 0,29-0,86	-	Mannequin PPMCC: 0,95-1 mean difference: 0.71-1.98 mm Students mean difference: 0.5-4 mm	-
	Digital analysis of digital photographs without categorical system	Dunk (62)	-	-	CoV: 17.08-4684.49	One session: ICC: 0.00-0.87 Three sessions: ICC_{mean}: 0.13-0.69	-	-	-
	Digital analysis of digital photographs without categorical system	Dunk (63)	-	-	-	Sagital view: ICC: 0.64-0.84 Deviation from vertical reference: ICC: 0.27-0.57 Posterior view: ICC: 0.13-0.69 Biological reference method: ICC 0.16-0.61 Vertical reference method ICC 0.13-0.69	-	-	-
	Digital analysis of digital photographs without categorical system • PAS	Ferreira (8)	ICC_{mean}: 68.4 ICC: 41.4-89.2	ICC_{mean}: 0.8 ICC: 0.21-0.97	-	-	-	-	-
	Digital analysis of digital photographs without categorical system	Fortin (65)	DC: 0.73-0.95	DC: 0.67-0.996	$SEM_{intra-rater}$: 0.5-.3.0^0 1.2-5.0 mm $SEM_{inter-rater}$: 0.5-2.9^0 1.6-5.8 mm	LoA inter sessions: -0.6-0.6 LoA inter-raters: -3.3-1.4	-	-	-

Supplement table 6 continued

	Instrument or observation method	Author (ref.)	Intra-rater reliability	Inter-rater reliability	SEM	Test-retest	Construct validity	Criterion/ Concurrent validity	Content validity
Digital assessment	Digital analysis of digital photographs without categorical system • DIPA	Furlanetto (66)	0.54-0.94	0.59-0.94	SEM: ≤ 3.6^0/≤ 6 mm MDC ≤ 6.9^0/≤ 11 mm	0.47-0.95	-	0.71-0.84 / p<0.001 correlation with X-ray	-
	Digital analysis of analog photographs (negatives) without categorical system	McEvoy (74)	-	-	-	ICC: 0.93-0.99	-	-	-
	Digital analysis of digital photographs without categorical system • PPAM	van Niekerk (77)	-	-	SEM: 3.33-11.09 mm	ICC: 0.94-0.99	-	--	-
	Digital analysis of digital photographs without categorical system • PosturePrint	Normand (78)	ICC: 0.64-0.88	ICC: 0.51-0.80	SEM: 0.7^0-2.7^0 2.6 -5.9 mm Absolute mean difference Interobservers: ≤ 8.4mm or ≤3.5^0 Intraobserver: ≤7.1mm or ≤3.2^0	-	-	-	-
	Digital analysis of digital photographs without categorical system	Paul (80)	-	-	-	-	-	Perspective error mean difference 3D-2D: 0.02^0-2.07^0	-
	Digital analysis of digital photographs without categorical system • PIAS	Pausic (81)	PIAS: ICC: 0.81-0.92	-	PIAS: SEM:0.55-0.88	-	Inter method agreement (PIAS – UTHSCSAIT) ICC: 0.992-0.998	-	-
	Digital analysis of digital photographs without categorical system	Perry (82)	Standing: ICC 0.868-1.00 Sitting: ICC 0.759-1.00 Slump: ICC 0.885-1.00	Standing ICC: 0.25-0.70 Sitting ICC: 0.24-0.73 Slump ICC: 0.19-0.54	Digitization SEM: 0.20^0-0.42^0 0.8-3.5mm Inter-rater SEM: 2.6^0-10^0 9.4 - 22.6mm	-	-	-	-

Supplement table 6 continued

	Instrument or observation method	Author (ref.)	Intra-rater reliability	Inter-rater reliability	SEM	Test-retest	Construct validity	Criterion/ Concurrent validity	Content validity
Digital assessment	Digital analysis of digital video stills without categorical system	Pownal (83)	-	-	SEM: 0.64^0-1.7^0 0 mm	Standing ICC: 0.35-0.94 CoV: 0.0-31.2% sitting ICC: 0.42-0.93 CoV: 0.6-83.3%	-	-	-
	Digital analysis of digital photographs without categorical system • PAS/SAPO	Ruivo (85)	0.66-0.83	0.87-0.96	SEM_{intra}= 2.72-4.03 mm SEM_{inter}= 1.64-2.35 mm	-	-	-	-
	Digital analysis of digital photographs without categorical system	Sanchez (86)	-	0.92-0.99	-	-	-	1.61^0-3.76^0 difference with VICON	-

BI= Barthel Index; CoV= coefficient of variation; GMFCS= Gross Motor Function Classification System; MAS=Modified Ashworth Score; PoA= percentage of Agreement ; VDS= Visual Descriptor Scale (for pain).

Supplement table 7

Supplement table 7: Table with the clinical usability values of the assessment instruments.

	Assessment category	Battery	Author (ref nr.)	Readability instructions	Time to administer	Costs of instrument	Physical needs	Interpretability of results	Patient load	Education time for observer	Total score
Body measurement : Direct method	Measurements with digital optic sensors of markers	MIDAS	McAlpine (73)	1	-1	0	1	-1	1	-1	**0**
	Manual analysis with ruler, goniometer & reference line	-	Harrison (67)	2	-1	2	2	1	1	1	**8**
Body measurement : Indirect method	Manual analysis of digital photographs	PAS	Pausic (81)	1	-1	0	1	0	1	-1	**1**
	Analog analysis of analog photographs	-	Zonnenberg (88)	2	-1	2	1	-1	1	1	**5**
Visual assessment: Direct method	Instantaneous visual analysis with pre-defined posture category system	-	Blanken (7)	2	-1	2	2	1	1	1	**8**
		CBE	Bunkan (58)	2	1	2	2	1	1	1	**10**
		IPR & MPR	Carr (59)	2	1	2	2	1	1	2	**11**
		RPE	Eriksson (64)	1	1	2	2	1	1	1	**9**
		SMT	Haugstad (68)	2	-1	0	2	0	1	1	**5**
		MS	Karim (70)	2	1	2	2	1	1	1	**10**
		GBE	Kvale (71)	2	-1	2	1	1	1	1	**7**
		REBA	Motamedzade (76)	2	1	2	2	2	1	1	**11**
		QEC	Motamedzade (76)	2	2	2	2	2	1	2	**13**
		PATH & sPATH	Paquet (79)	1	2	2	2	1	1	1	**10**
		mRULA	Menendez (75)	2	-1	2	2	1	1	1	**8**
		UCCUC	Menendez (75)	2	-1	2	2	1	1	1	**8**
Visual assessment: indirect method - photograph	Visual analysis of photographs with pre-defined posture category system	-	Barks (55)	2	1	2	2	1	1	2	**11**
		QWAS	de Bruijn (57)	2	2	2	2	1	1	2	**12**
		-	Fedorak (29)	0	2	2	2	2	1	1	**10**
		REBA	Hignett (69)	2	1	2	2	2	1	1	**11**
		-	Liebregts (72)	1	0	1	2	1	1	1	**7**
		PPAS	Rodby-Bousquet (84)	0	-1	2	2	1	1	1	**6**
		-	Wilbanks (87)	0	1	1	2	0	1	1	**6**
Visual assessment: indirect method - video	Visual analysis of video with pre-defined posture category system	VAS	Ackermann (21)	1	1	1	1	1	1	1	**7**
		mRULA	Bao (52)	1	-1	0	1	1	1	-1	**2**
		-	Bao (53)	1	-1	0	1	1	1	0	**3**
		QEC	David (60)	2	2	2	2	2	1	2	**13**
		RULA	Dockrell (61)	2	1	2	2	2	1	1	**11**
Digital Assessment	Digital analysis of analog photographs without categorical system	Videopoint	Barker (54)	0	-1	1	1	1	1	1	**4**
		3D-PAT	Brink (56)	1	-1	-2	1	-1	1	-1	**-2**
		-	Dunk (62)	0	2	0	1	1	1	1	**6**
		-	Dunk (63)	0	-1	0	1	-1	1	-1	**-1**
		PAS	Ferreira (8)	0	-1	0	1	-1	1	-1	**-1**
		-	Fortin (65)	1	-1	0	-1	-1	1	-1	**-2**
			Furlanetto (66)	2	-1	-1	1	1	1	1	**4**
		-	McEvoy (74)	1	-1	0	1	-1	1	1	**2**
		PPAM	van Niekerk (77)	1	-1	-1	1	-1	1	1	**1**
		Postureprint	Normand (78)	2	-1	-1	1	-1	1	1	**2**
		-	Paul (80)	2	2	2	2	1	1	1	**11**
		PIAS	Pausic (81)	1	-1	0	1	0	1	-1	**1**
		-	Perry (82)	1	-1	-1	1	-1	1	-1	**-1**
		-	Pownall (83)	2	-1	-1	1	-1	1	1	**2**
		-	Ruivo (85)	1	-1	1	1	-1	-1	1	**1**
		-	Sanchez (86)	1	-1	0	1	-1	-1	-1	**-2**

Supplement table 8

Supplement table 8: Probability of publication bias.

Category of Correlation coefficient[1]	Small studies ($N_{population} < 30$)	Large studies ($N_{population} \geq 30$)
≥ 75%[2]: ≥ 0.75	7-8[4]	2
≥ 75%[2]: 0.50	8	5-6[4]
Others[3]	1	3-4[4]

[1] The values: ≥ 0.75: good reliability, 0.50 - 0.75: moderate reliability, < 0.5: poor reliability. Three categories (according to GRADE's approach):

[2] if ≥ 75% of correlations coefficients ≥ 0.75 respectively 0.5/study; [3] if not applicable to the first two categories (see text: 'interpretation of results').[4] If not clear if all criteria are applied to all studies in one of the categories, the number of studies is indicated with >1 number. Correlation coefficients are based on the inter-rater values (or test-retest value if no inter-rater value is available) per studies.